

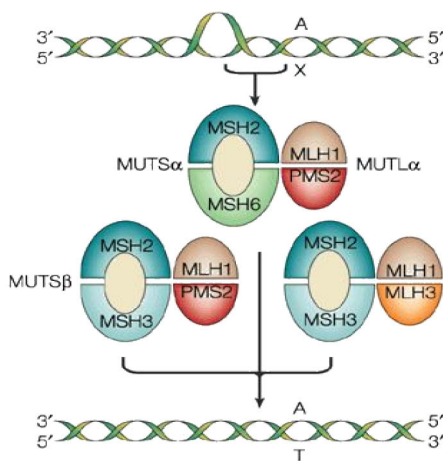


**UNIVERSITA' DI NAPOLI FEDERICO II**

**DOTTORATO DI RICERCA IN BIOCHIMICA  
E BIOLOGIA CELLULARE E MOLECOLARE  
XXVII CICLO**

**Raffaella Liccardo**

**THE MULTIVARIATE ANALYSIS OF UNKNOWN  
SIGNIFICANCE VARIANTS IN MISMATCH  
REPAIR GENES DEMONSTRATES THEIR  
PATHOGENICITY IN LYNCH SYNDROME**



**Academic Year 2013/2014**



**UNIVERSITA' DI NAPOLI FEDERICO II**

**DOTTORATO DI RICERCA  
BIOCHIMICA E BIOLOGIA CELLULARE E MOLECOLARE  
XXVII CICLO**

**THE MULTIVARIATE ANALYSIS OF UNKNOWN  
SIGNIFICANCE VARIANTS IN MISMATCH  
REPAIR GENES DEMONSTRATES THEIR  
PATHOGENICITY IN LYNCH SYNDROME**

**Candidate  
Raffaella Liccardo**

Tutor  
Prof. Paola Izzo

Coordinator  
Prof. Paolo Arcari

Co-Tutor  
Dr. Francesca Duraturo

**Academic Year 2013/2014**

## RIASSUNTO

Il cancro ereditario non poliposico del colon-retto, noto anche come Sindrome di Lynch, rende conto di circa il 3-5% dei tumori coloretali sporadici; è una sindrome ereditaria autosomica dominante, associata a mutazioni germinali nel complesso dei geni del riparo di appaiamenti errati di basi del DNA (Mismatch Repair, MMR). La perdita di funzione di una o più di queste proteine a livello germinale determina una significativa instabilità genomica a livello somatico, specie a carico di sequenze ripetute (microsatelliti) presenti in molti oncogeni e oncosoppressori, accelerando così il processo di tumorigenesi. Ciò si traduce in una più precoce età di insorgenza della malattia (circa 45 anni) rispetto ai casi sporadici di cancro del colon. Tale sindrome presenta una predisposizione al tumore non solo del colon, ma anche di altri organi quali endometrio, stomaco, ovaie, piccolo intestino, epitelio epatobiliare, epitelio uroepiteliale e cervello. Alla base di questa variabilità fenotipica esisterebbe un'interazione tra geni principali (MMR) e geni modificatori e/o fattori ambientali.

I geni MMR maggiormente coinvolti sono MLH1 e MSH2, che risultano mutati rispettivamente nel 40% e 39% dei casi HNPCC, mentre mutazioni nei geni MMR cosiddetti "minori", MSH6, PMS2, MLH3 e MSH3, rendono conto complessivamente del 21% dei casi, con un maggiore contributo del gene MSH6 (11%).

L'iter diagnostico comunemente utilizzato prevede la ricerca di mutazioni puntiformi, mediante DHPLC e diretto sequenziamento, e di ampi riarrangiamenti, mediante MLPA, nei geni principali MLH1 e MSH2, per la caratterizzazione molecolare di soggetti con diagnosi clinica di Sindrome di Lynch. Tuttavia, tale indagine non sempre fornisce risultati informativi ai fini della consulenza genetica. Infatti, nel nostro laboratorio, negli ultimi 5 anni, su 117 famiglie con predisposizione ereditaria allo sviluppo del tumore, selezionate in base a criteri diagnostici internazionali (*Criteri di Amsterdam, Lynch HT et al, ICG-HNPCC, Gastroenterology 1999, e Linee Guida di Bethesda, Umar et al, J Natl Cancer Inst., 2004*), 64 sono risultate senza una chiara diagnosi molecolare, per assenza di mutazioni nei due geni principalmente investigati, o per la presenza in tali geni di varianti di dubbio significato patogenetico (VUS, Variants of Uncertain Significance) (varianti missense, introniche e silenti). I soggetti appartenenti a tali famiglie, di cui 40 rispondenti ai Criteri di

Amsterdam e 24 alle Linee Guida di Bethesda, sono stati selezionati per più estese indagini molecolari.

Al fine di valutare il significato patogenetico di varianti VUS, sono stati utilizzati diversi approcci sperimentali quali analisi di segregazione, studi di popolazione, valutazione di instabilità genomica (MSI, instabilità di sequenze Microsatelliti) nel tessuto tumorale, identificazione di perdita di eterozigosità (LOH) nel tessuto tumorale, analisi quantitativa di trascritti e di prodotti proteici ed infine analisi bioinformatiche. La combinazione di più risultati concordanti tra di loro ha facilitato l'interpretazione del significato patogenetico delle varianti analizzate, secondo le linee guida riportate in letteratura (*Colon Cancer Family Registry 2009, InSiGHT Variant Interpretation Committee 2011*).

In particolare, è stato approfondito lo studio di quattro VUS, di cui tre identificate nel gene MLH1 e una nel gene MSH2, in quanto associate a fenotipi clinici piuttosto severi, e spesso segreganti con la malattia nelle famiglie portatrici di tali varianti.

Molto interessante è risultato lo studio delle due mutazioni identificate nelle regioni 3'UTR dei geni MMR (c\*30\_32delTTC nel gene MLH1 e c\*226A>G nel gene MSH2). In particolare, la c\*30\_32delTTC, riportata come variante benigna nel database dei geni MMR (*Insight-group database*), sulla base dei nostri risultati, è risultata determinare un abbassamento nei livelli sia del trascritto genico che del suo prodotto proteico. I risultati ottenuti sono in accordo con recenti dati di letteratura in base ai quali la regione in cui cade la mutazione è stata identificata quale putativo sito target del miRNA miR-422a (*Mao et al., J Biol Chem, 2008*). Più recentemente, è stato anche dimostrato un meccanismo di regolazione a feedback tra MLH1 e miR-422a (*Mao et al., Cell Res, 2012*). È interessante notare che tale mutazione, nei casi di fenotipo più aggressivo, è stata da noi identificata sempre in associazione con un'altra variante, la c.454-51t>c. Pertanto, abbiamo ipotizzato anche un probabile effetto additivo patogenetico tra queste due alterazioni geniche.

La variante nel gene MSH2, la c\*226A>G, anch'essa già descritta in letteratura e riportata come variante benigna, nella famiglia da noi identificata è risultata associata con la malattia. Inoltre, in questa famiglia abbiamo osservato l'associazione anche con condizioni morbose diverse dalla Sindrome di Lynch, come il Linfoma di Hodgkins. I nostri risultati hanno mostrato un incremento nei livelli

sia del trascritto che della proteina. In seguito a tali risultati, al fine di chiarire il meccanismo molecolare alla base dell'iper-espressione da noi osservata, è stata effettuata un'analisi in silico mediante i programmi di predizione di zone target di miRNA (TargetScan e miRanda) e di fattori di regolazione trascrizionale (TRANSFAC). La regione in cui cade la mutazione è stata identificata quale putativo sito target di due miRNA (hsa-miR-137, hsa-miR47953p), e di quattro fattori di regolazione trascrizionale, ZNF333, POU6F1, CDP e PMX1, con principale ruolo di repressori trascrizionali. Pertanto, abbiamo ipotizzato che tale variante, impedendo il legame della regione 3'UTR di MSH2 con tali fattori, possa determinare una incontrollata espressione del trascritto e del relativo prodotto proteico. I risultati ottenuti sono stati da noi confermati mediante saggio funzionale di espressione di luciferasi che hanno mostrato un aumento dei livelli di espressione in presenza di tale variante, supportando quindi alcuni dati di letteratura che mostrano come anche l'iper-produzione delle proteine del mismatch repair possa essere considerata deleteria (Zhang H. et al., *Cancer Res* 1999; Shcherbakova PV. et al., *Mol. Cell. Biol.*, 2001). I risultati ottenuti per la variante c\*226A>G nel gene MSH2 sono a favore di un possibile ruolo patogenetico nello sviluppo della malattia.

L'attività di ricerca è proseguita con l'identificazione di varianti nei geni minori del mismatch repair, per quei pazienti risultati negativi per qualsiasi tipo di alterazione nei geni maggiori, MLH1 e MSH2. Infatti, alla luce di recenti studi che hanno dimostrato come il sistema MMR, oltre alla riparazione post-replicativa, svolge altri ruoli altamente rilevanti nella carcinogenesi, come il controllo del ciclo cellulare e del processo apoptotico (Jiricny J. et al. *Mol Cell Biol* 2006, Ji et al., *BMC Med* 2012, Mao G., et al., *Cell Res*, 2012), è ipotizzabile che in questi nuovi ruoli siano coinvolti anche i geni MMR minori. Oltre a varianti di certo significato patogenetico (due mutazioni frameshift ed un'inserzione in frame nel gene MSH6, un'ampia duplicazione nel gene PMS2 ed una mutazione troncante nel gene MLH3) identificate in questo studio, e varianti la cui patogenicità è stata già descritta in letteratura (cinque mutazioni missense nel gene MSH6), un cospicuo numero di VUS, circa 70, sono state ritrovate anche nei geni minori. Pertanto, una dettagliata caratterizzazione fenotipica, come descritto sopra per i geni MLH1 e MSH2, è stata effettuata anche per alcune di tali varianti, allo scopo di classificarle come patogenetiche. E'

risultato interessante notare come molti tra i soggetti analizzati presentano multiple alterazioni geniche in più geni MMR, a sostegno dell'ipotesi di un'ereditarietà complessa, di tipo poligenico, per la Sindrome di Lynch (*Durauto F. et al. Int J Cancer 2011*) che è stata sollevata sulla base di recenti studi su genoma di lievito (*Martinez SL. et al. PNAS, 2010, Kumar C. et al. Mol Biol, 2011*). Purtroppo, per la scarsa collaborazione dei soggetti analizzati, solo per una delle 64 famiglie studiate è stato possibile verificare in maniera completa come l'associazione simultanea di più varianti VUS identificate sia nei geni MMR major che in quelli minor, segregava perfettamente con la malattia. L'ipotesi di un effetto sinergico tra alleli a bassa penetranza spiegherebbe anche la mancata segregazione osservata per le due varianti di certo significato patogenetico identificate nel gene MSH6, c.3261\_62insC e c.3296\_97 delTT. Infatti, solo in alcuni dei membri affetti delle famiglie analizzate queste due mutazioni sono state trovate associate a varianti negli altri geni minori. Inoltre, non si esclude la concomitante presenza di alterazioni in altri geni pur implicati nel processo di carcinogenesi, non investigati in tale studio. In conclusione, l'attività scientifica svolta ha consentito in primo luogo di chiarire meglio le correlazioni genotipo-fenotipo nella Sindrome di Lynch, dimostrando l'importanza di un approccio analitico multidisciplinare, anche al fine di comprendere i meccanismi molecolari alla base di fenotipi-malattia non canonici; in secondo luogo, sono state messe in luce nuove prospettive diagnostiche per la Sindrome di Lynch, che si discostano dalla classica trasmissione monogenica, a favore dell'ipotesi di un modello di ereditarietà di tipo poligenico, anche se ulteriori studi sono necessari per avvalorare tale ipotesi.

## SUMMARY

The Hereditary Non-Polyposis Colorectal Cancer (HNPCC), also known as Lynch Syndrome (LS), is an autosomal dominantly inherited cancer syndrome that accounts for about 3-5% of all colorectal cancers (CRCs). It is commonly associated with germline mutations in the mismatch repair (MMR) genes. The loss of function of one or more of these proteins results in a significant genomic instability at somatic level, particularly in repetitive DNA sequences (microsatellites) present in many oncogenes and tumor suppressor genes. This mutator phenotype promotes the tumorigenesis process that justifies the earlier age of onset of the disease (approximately 45 years). LS is characterized by high lifetime risk for tumor development, especially CRC, endometrial cancer and other extracolonic tumors. These extra-colonic malignancies include carcinomas of the small intestine, stomach, pancreas, biliary tract, ovarium, upper urinary tract and brain.

The MLH1 and MSH2 mutations account for about 40% and 39% of HNPCC cases, respectively, while mutations in minor MMR genes, MSH6, PMS2, MLH3 and MSH3, justify a total of 21% of cases, in which the MSH6 gene gives a greater contribution (11%).

The molecular characterization of patients with a clinical diagnosis of Lynch Syndrome relies on the identification of point mutations by DHPLC and direct sequencing, and large rearrangements by MLPA, in the major MMR genes, MLH1 and MSH2.

This strategy does not always provide exhaustive information for genetic counseling. Indeed, we analyzed 117 families selected according to international diagnostic criteria (*Amsterdam Criteria*, *Lynch HT et al*, *ICG-HNPCC*, *Gastroenterology* 1999 and *Bethesda Guidelines*, *Umar et al*, *J Natl Cancer Inst.*, 2004). Molecular diagnosis was achieved for 53 families while no MLH1/MSH2 mutations were identified in the remaining 64 families. Moreover, we identified genetic variants of uncertain significance (VUS) (missense, intronic and silent variants) in several other patients.

Therefore, in this study we have characterized the VUS that we had identified in the major MMR genes, in order to classify them as pathogenic, using the following approaches: segregation analysis, population studies to exclude the polymorphic nature of the variant, assessment of Microsatellite Instability (MSI), gene expression studies both at mRNA and protein levels and *in silico* analysis with a variety

of bioinformatics tools such as HSF (Human splicing Finder), PolyPhen (Polymorphism Phenotyping), SIFT (Sorting Intolerant From Tolerant), and PredictProtein.

According to literature data (*Colon Cancer Family Registry in 2009, Insight Variant Interpretation Committee 2011*), a combination of these strategies has been used for each variant in order to assess the pathogenicity of uncertain variants.

We performed phenotypic and functional characterization of some of the most interesting variants of uncertain significance identified (three variants in the MLH1 gene and one in the MSH2 gene) that were associated with severe disease phenotypes.

In particular, the study of two mutations detected in the 3'untranslated regions (3'UTR) of the MLH1 and MSH2 genes (c\*30\_32delTTC in the MLH1 gene and c\*226A>G in the MSH2 gene) gave very interesting results.

The c\*30\_32delTTC was reported as a benign variant in the MMR gene variants database (*Insight-group database*). Our results showed lower levels both at RNA and protein levels. These data are in agreement with recent literature studies reported on this variant (*Mao G. et al., J Biol Chem, 2008*). In particular, the region in which falls the mutation has been identified as a putative target point of the miR-422a and a mechanism of feedback regulation between MLH1 and this miRNA was also demonstrated (*Mao G. et al., Cell Res, 2012*). In our study, in more aggressive phenotypes, this mutation has always been identified in association with another variant, the c.454-51t> c. Therefore, a likely pathogenetic additive effect between these two genetic alterations could be proposed.

The multivariate analysis of the other 3'UTR variant, the c.\*226A>G in MSH2 gene, also provided very interesting data in order to assess the correlation with the disease phenotype. In our study, this variant was found associated with both typical Lynch Syndrome features (colorectal and endometrial tumors) and atypical phenotypes such as Hodgkin Lymphoma. This analysis showed increased mRNA and protein levels, as also confirmed by a functional luciferase assay. In order to clarify the molecular mechanism at the basis of this over-expression, *in silico* analysis was performed for prediction of miRNA target sites (TargetScan and MiRanda) and transcriptional regulation factor binding sites (TRANSFAC). The region in which falls the mutation is identified as a putative target point of two miRNAs (hsa-



miR-137, hsa-miR47953p), and four trans-acting protein factors, ZNF333, POU6F1, CDP and PMX1, known also as transcriptional repressors. Therefore, we hypothesized that this variant could prevent the binding of these factors with the MSH2 3'UTR leading to unregulated expression of the MSH2 gene. In agreement with several literature data showing a deleterious effect derived from overproduction of MMR proteins (*H. Zhang et al., Canc Res 1999; Shcherbakova PV. et al., Mol. Cell. Biol., 2001*), it is conceivable that the variant c.\*226A> G in the MSH2 gene has a pathogenic role in the development of the disease.

In our study we also analyzed the minor MMR genes, MSH6, PMS2, MLH3 and MSH3, for the presence of germline variants in patients who had been resulted to be negative for germline mutations in the major MMR genes. Recently, literature data suggest that the MMR proteins, including the minor MMR factors, may have other functions in cell cycle and apoptosis regulation, in addition to the post-replicative repair role, that could be highly relevant for the carcinogenesis process (*Jiricny J. et al. Mol Cell Biol 2006, Ji et al., BMC Med 2012, Mao G., et al., Cell Res, 2012*). According to these data, we undertook a mutational analysis on the minor MMR genes and were able to identify 77 variants, five of which were pathogenic mutations, 5 missense variants already described in the literature as pathogenic mutations and 67 unclassified variants (missense, silent and intronic variants). A detailed phenotypic characterization, as described above for the MLH1 and MSH2, was also carried out for some of these variants, in order to classify them as likely pathogenetic. The most relevant result of this mutational analysis on minor MMR genes was the simultaneous presence of multiple molecular alterations in different genes or in a single gene in several unrelated patients. We proposed that some (or all) of these variants could constitute low penetrance alleles, with an additive effect on the risk of the disease (*Durauto F. et al., Int J Cancer 2011*). This hypothesis is confirmed by an exhaustive molecular analysis performed for one of the 64 subjects analyzed in this study. Two literature studies on the yeast genome also demonstrated our hypothesis (*Martinez SL. et al. PNAS, 2010, Kumar C. et al. Mol Biol, 2011*). The probability of a synergistic effect between low-penetrance alleles could also explain the loss of segregation observed for two pathogenetic variants identified in the MSH6 gene, c.3261\_62insC

and c.3296\_97 delTT. These two mutations have been found associated with variants in other minor MMR genes in some of affected members of the families analyzed. Moreover, we can not exclude the concomitant presence of alterations in other genes involved in carcinogenesis but not investigated in this study.

In conclusion, this study allowed to clarify the genotype-phenotype correlations in Lynch syndrome, demonstrating the importance of a multifactorial likelihood analytical approach, also in order to understand the molecular mechanisms that regulate the MMR protein and their putative prognostic and therapeutic implications in Lynch Syndrome. Moreover, the simultaneous presence of molecular alterations in several genes (major and minor MMR genes) could suggest an additive effect of these mutations in cancer predisposition, against the classical monogenic transmission and in favor of a polygenic inheritance, although further studies on other familial cases will be needed to confirm this hypothesis.

# INDEX

	Pag.
<b>1. INTRODUCTION</b>	<b>1</b>
1.1 Inherited colorectal cancer syndromes	1
1.2 Lynch Syndrome	4
1.3 New insights into the molecular features of Lynch Syndrome	11
1.4 Genotype-phenotype associations in Lynch Syndrome:	
Canonical features	12
Non-canonical features	14
1.5 Characterization of the "variants of uncertain significance" (VUS) in the MMR genes	17
1.6 Probability of a "synergistic effect" between low risk allelic variants in the MMR genes	19
1.7 Scientific hypothesis and aim of the work	21
<b>2 MATERIALS AND METHODS</b>	<b>23</b>
2.1 Patients	23
2.2 DNA, RNA and protein isolation	23
2.3 Microsatellite Instability (MSI) analysis	24
2.4 Immunohistochemistry analysis	25
2.5 MMR germline point mutations and large genomic rearrangements analysis	26
2.6 RNA-based analysis	27
2.7 Protein analysis	28
2.8 In silico analysis	29
2.9 Functional analysis	31
<b>3. RESULTS</b>	<b>33</b>
3.1 Phenotypic and functional characterization of VUS in the Major MMR genes	34
3.2 Mutation detection analysis of minor MMR genes	50
<b>4. DISCUSSION/CONCLUSIONS</b>	<b>57</b>
<b>5. REFERENCES</b>	<b>64</b>

## 4 LIST OF TABLES AND FIGURES

	<b>Pag.</b>
<b>Table 1.</b> Hereditary colon cancer syndromes: clinical and genetic features	3
<b>Table 2.</b> Clinical-pathological features of Lynch syndrome	4
<b>Table 3.</b> The Amsterdam Criteria	10
<b>Table 4.</b> The Bethesda Guidelines	10
<b>Table 5.</b> Proposed classification system for MMR variant interpretation	18
<b>Table 6.</b> Microsatellites analysis with the Gene Quality CC-MSI kit (AB-Analytical).	24
<b>Table 7.</b> Primer sequences and amplicon sizes for MLH1 and MSH2 mRNA quantification.	29
<b>Table 8.</b> VUS identified in the MLH1 gene in patients with MSI-H.	34
<b>Table 9.</b> Families with the c.*30_32delTTC and c.454-51t>c mutations in the MLH1 gene.	40
<b>Table 10.</b> VUS identified in the MSH2 gene in patients with MSI-H and MSI-L	45
<b>Table 11.</b> Patients with more mutations in the MMR genes.	53
 <b>Figure 1.</b> Distribution of the different types of colorectal cancer	 1
<b>Figure 2.</b> The human Mismatch repair system	5
<b>Figure 3.</b> Two molecular pathways can lead to CRC with MSI	8
<b>Figure 4.</b> Graphic representation of the PSICHECK2 vector.	32
<b>Figure 5.</b> Identification of the c.304G>A mutation in the MLH1 gene	34
<b>Figure 6.</b> In silico analysis of the c.304G>A mutation identified in the MLH1 gene.	35
<b>Figure 7.</b> Analysis of MLH1 cDNA in a patient carrier of the c.304G>A variant.	36
<b>Figure 8.</b> HSF-ESEFINDER analysis of the MLH1 c.304G>A mutation and sequencing analysis of the MLH1 cDNA ex2-4 abnormal fragment	37
<b>Figure 9.</b> Identification of the c.438A>G mutation in the MLH1 gene.	37

<b>Figure 10.</b> HSF-ESEFINDER analysis for the MLH1 c.438A>G mutation.	<b>37</b>
<b>Figure 11.</b> Polyacrylamide electrophoresis analysis of the MLH1 cDNA in a patient carrier of the variant c.438 A>G.	<b>38</b>
<b>Figure 12.</b> q-Real-Time PCR analysis of the MLH1 mRNA in a patient with the c.438A>G variant.	<b>39</b>
<b>Figure 13.</b> Identification of the c.*30_32delTTC mutation in the MLH1 gene.	<b>39</b>
<b>Figure 14.</b> HSF analysis of the c.454-51t>c variant.	<b>40</b>
<b>Figure 15.</b> TRANSFAC analysis of the MLH1 3'UTR.	<b>41</b>
<b>Figure 16.</b> MLH1 cDNA q-Real-Time PCR analysis in a patient with the two variants: c.*30_32delTTC and c.454-51t>c.	<b>42</b>
<b>Figure 17.</b> Western blot analysis of MLH1 expression in a Patient with two variants: c.*30_32delTTC and c.454-51t>c.	<b>42</b>
<b>Figure 18.</b> MLH1, MSH2, MSH6 immunohistochemistry (IHC) in a patient carrier of the c.*30_32delTTC and c.454-51t>c mutations.	<b>43</b>
<b>Figure 19.</b> Identification of the c.*226A>G mutation in the MSH2 gene.	<b>44</b>
<b>Figure 20.</b> Pedigree with the segregation of the c.*226A>G mutation.	<b>44</b>
<b>Figure 21.</b> TRANSFAC analysis of the MSH2 3'UTR.	<b>45</b>
<b>Figure 22.</b> In silico analysis of the MSH2 3'UTR by the prediction tools TargetScan and miRanda.	<b>47</b>
<b>Figure 23.</b> q-Real-Time PCR analysis of the MSH2 mRNA in the patients with the c.*226A>G variant.	<b>48</b>
<b>Figure 24.</b> Western blot analysis of MSH2 expression in the index patient with the c.*226A>G variant.	<b>48</b>
<b>Figure 25.</b> MSH2 immunohistochemistry (IHC) results in the III-2 index patient.	<b>49</b>
<b>Figure 26.</b> MSH2 immunohistochemistry (IHC) results in the II-6 patient.	<b>49</b>
<b>Figure 27.</b> Schematic diagram of WT and Mut luciferase reporter gene constructs.	<b>49</b>
<b>Figure 28.</b> Relative Luciferase expression.	<b>50</b>

<b>Figure 29.</b> q-Real-Time PCR analysis of endogenous MSH2 mRNA in the HT29 and SW480 cells transfected with miR-137 and negative control (FAM).	<b>50</b>
<b>Figure 30.</b> Pathogenic mutations identified in minor MMR genes.	<b>51</b>
<b>Figure 31.</b> Histogram of several VUS identified in the minor MMR genes.	<b>51</b>
<b>Figure 32.</b> (A) Pedigree of a family with simultaneous MMR gene variants.	
(B) Genotypes identified in the patients of the family.	<b>54</b>
<b>Figure 33.</b> Microsatellite instability analysis for the II-5 family subject.	<b>54</b>
<b>Figure 34.</b> MSH6, MSH2 and MLH1 immunohistochemistry (IHC) in the patient carrier of the c.3261_62insC variant in MSH6 exon 5.	<b>55</b>
<b>Figure 35.</b> MLH1, MSH2 and MSH6 immunohistochemistry (IHC) in the patient carrier of the c. 2049_2050insAGT variant in MSH6 exon 4.	<b>56</b>
<b>Figure 36.</b> Model for reciprocal feedback regulation between MLH1 and miR-422a	<b>61</b>
<b>Figure 37.</b> Predicted folding of MLH1 WT and MT 3'-UTR	<b>62</b>

## 1. INTRODUCTION

### 1.1 Inherited colorectal cancer syndromes

Colorectal cancer (CRC) is a multifactorial disease in which genetic and environmental factors are involved.

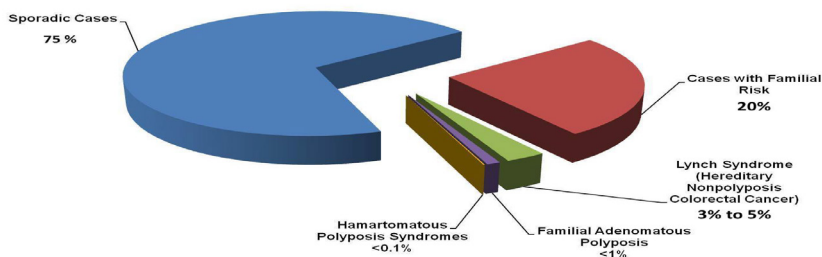
CRC is extremely common as reflected by its worldwide annual incidence of  $1.2 \times 10^6$  cases. Familial CRC, in which one or more first-degree and/or second degree relatives of the index case manifest CRC, constitutes approximately 20% of the total CRC burden (248,000 cases worldwide) (Lynch H.T., 2014).

High penetrance mutations confer a predisposition to CRC in the so-called hereditary syndromes, responsible for about 2-6% of the total CRC. Low-penetrance mutations are found in the remaining part of CRC (about 96%), representing a risk factor in both sporadic and familial cases (Chang CC., 2011, Valle L, 2014).

CRC syndromes are defined on the basis of clinical, pathological and, more recently, genetic findings (Valle L., 2014) (Tab. 1).

Accordingly, the identification of predisposing genes allows for accurate risk assessment and more precise screening approaches.

Lynch syndrome (LS) is by far the most common hereditary form of CRC with an incidence of 3-5% of all CRCs whereas its primary genetic counterpart, namely Familial Adenomatous Polyposis (FAP), accounts for less than 1% of the total CRC burden (Lynch H.T., 2014) (Fig.1).



**Figure 1.** Distribution of the different types of colorectal cancer (Lynch et al., Curr Treat Option Oncol, 2014)

HNPCC and FAP are diseases with autosomal dominant inheritance, caused by germline mutations in the DNA Mismatch Repair genes (MMR), or in the Adenomatous Polyposis Coli tumor suppressor gene (APC), respectively. These syndromes may also occur in more attenuated forms. In FAP syndrome, attenuated forms (AFAP) are caused by low penetrance mutations (missense mutations) in the main APC gene or by biallelic loss of the MYH gene (MAP, MUTYH-associated polyposis with autosomal recessive inheritance), encoding a protein of the Base Excision Repair complex (BER).

Variant forms of HNPCC are characterized by the presence of additional tumors in extra-colonic locations of still unclear etiology. Recent studies suggest that an interaction between main genes (MMR) and modifier genes and / or environmental factors may be at the basis of these tumors. These variant syndromes include Muir-Torre syndrome (autosomal dominant) due to MSH2 and MLH1 genes mutations and characterized by the presence of cutaneous manifestations (multiple sebaceous adenomas, epithelioma, keratoacanthoma) associated with colorectal and endometrial cancers; Turcot syndrome (autosomal dominant) associated with APC, PMS2 and MLH1 genes mutations, wherein brain cancers (glioblastoma and cerebellar medulloblastoma) are associated with colorectal cancer (*Sammader NJ., 2014*).

More recently, gastric cancers have been included in the tumor spectrum of HNPCC. The molecular and clinicopathological profiles of gastric cancers in HNPCC mutation carriers have been evaluated and compared with the profiles of sporadic gastric cancers, and several differences have been identified, while there were similarities with canonical HNPCC spectrum malignancies. Stomach can thus be considered as a target tissue where somatic inactivation ("second hit") of MMR genes may occur in carriers of a germline mutation ("first hit") (*Corso G., 2011; Gylling A., 2007*).

In the process of colorectal carcinogenesis many other genes are involved such as oncogenes and tumor suppressor genes that play a key role in the control of cell cycle. Mutations in these genes are at the basis of rarer inherited CRC syndromes. These are mainly "hamartomatous polyposis syndromes" characterized by the presence



## Introduction

of benign adenomas arising from epithelial and / or stromal intestinal tissue, which increase the risk of developing CRC. These syndromes, whose characteristics are summarized in Table 1, include Peutz-Jeghers syndrome, juvenile polyposis, Cowden syndrome and Bannayan-Riley-Ruvalcaba syndrome (*Jelsig AM., 2014, Stoffel EM., 2015*).

Disease (OMIM)	Gene	Incidence	Inheritance	Mutation identified (%)	Penetrance	Clinical features
Hereditary non polyposis colorectal cancer (HNPCC) (114500)	MLH1, MSH2, MSH6, PMS2, MLH3, EPCAM	1 in 400	AD	Point mutations, Large rearrangements (60-80%)	90%	Proximal CRC, endometrial carcinoma, ovarian tumors, small bowel carcinoma, urinary tract carcinoma
Classical familial adenomatous polyposis (FAP) (175100)	APC	1 in 8000	AD	Point mutations, Large rearrangements (80-90%)	<100%	100 to >500 adenomatous polyps of large bowel, duodenum, stomach
Attenuated FAP (AFAP) (175100)	APC	<1 in 8000	AD	Point mutations, Large rearrangements (20-30%)	<100%	10 to 100 adenomatous polyps of large bowel, duodenum, stomach
MUTYH-associated polyposis (MAP) (608456)	MUTYH	<1 in 10000	AR	Point mutations, Large rearrangements (15-20%)	<100%	20 to 100 adenomatous polyps of large bowel, duodenum, stomach
Muirre-Torre syndrome (HNPCC) (158320)	MLH1, MSH2	<1 in 400	AD	Point mutations, Large rearrangements (60-80%)	90%	CRC, endometrial carcinoma multiple sebaceous adenomas, epithelioma, keratoacanthoma
Turcot syndrome (HNPCC) (276300)	APC, PMS2, MLH1	<1 in 400	AD	Point mutations, Large rearrangements (60-80%)	90%	CRC, glioblastoma, cerebellar medulloblastoma
Peutz-Jeghers syndrome (PJS)(175200)	STK11 (LKB1)	1 in 200000	AD	Point mutations, Large rearrangements (90%)	95-100%	<20 Juvenil polyps (PJ) of large bowel, duodenum, stomach, mucocutaneous/perioral hyperpigmentation, ovarian tumors, breast cancer
Juvenile polyposis syndrome (JPS) (174900)	SMAD4, BMPR1A	1 in 100000	AD	Point mutations, Large rearrangements (60%)	90-100%	5 to 100 JP of large bowel, duodenum, stomach, gastric cancer
Cowden syndrome (CS) (158350)	PTEN	1 in 200000	AD	Point mutations, Large rearrangements (80%)	90-95%	Multiple JP/lipomas of large bowel, duodenum, stomach, mucocutaneous tumors, breast cancer, endometrial carcinoma, thyroid cancer
Bannayan-Ruvalcaba-Riley syndrome (BRRS) (153480)	PTEN	1 in 200000	AD	Point mutations, Large rearrangements (60%)	90-95%	Multiple JP/lipomas of large bowel, duodenum, stomach, macrocephaly, developmental delay, hemangiomas

**Table 1.** Hereditary colon cancer syndromes: clinical and genetic features (AD, autosomal dominant; AR, autosomal recessive).

## **1.2 Lynch Syndrome**

The characterization of families affected by hereditary colon cancer was described for the first time in 1913 by Warthin. In 1966, Lynch and his colleagues described two families who had colorectal cancer associated with endometrial and gastric cancers. This syndrome is known with the term of HNPCC (Hereditary Non-Polyposis Colorectal Cancer) to highlight the absence of colon polyps and distinguish this syndrome from other hereditary forms of colorectal cancer such as familial adenomatous polyposis (FAP) (*Peltomaki P., 2001; Weitz J., 2005; Olschwang S., 2006*).

Hereditary non-polyposis colorectal cancer (HNPCC), also known as Lynch syndrome (LS), accounts for about 3-5% of sporadic colorectal cancers (CRCs); it is an autosomal dominant condition with recessive phenotype caused by a defect in one of the mismatch repair (MMR) genes.

The main features of the Lynch syndrome are listed in Table 2 (*Lynch HT., 2014; val Lier MGF., 2010*).

---

<b>Autosomal dominant inheritance;</b>
<b>Penetrance for colorectal cancer (CRC) of 85-90%;</b>
<b>Earlier age of onset of CRC (~ 45 years ) with respect to general population (69 years);</b>
<b>Preferential tumor localization in the right-sided colon;</b>
<b>Presence of multiple synchronous and metachronous colorectal cancers;</b>
<b>Better prognosis than CRCs;</b>
<b>Increased risk for extra-colonic cancers;</b>
<b>Accelerated carcinogenesis;</b>
<b>Poorly differentiated tumors, with a marked lymphocytic peritumoral inflammation recalling features of the so-called "Crohn's reaction";</b>
<b>Microsatellite instability</b>

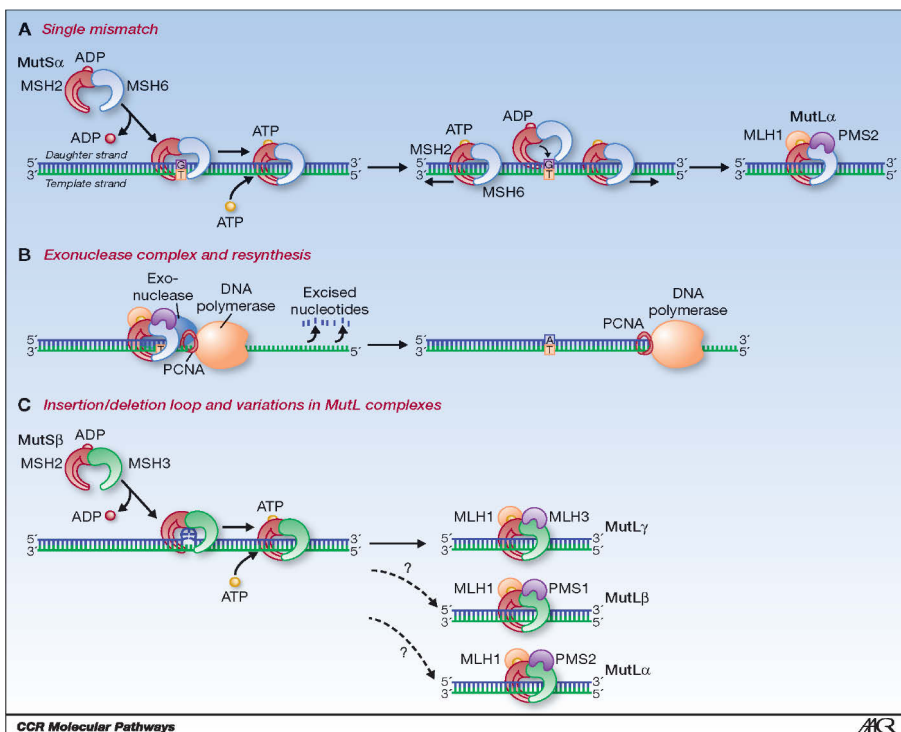
---

**Table 2.** Clinical-pathological features of Lynch syndrome.

LS patients present a germline mutation in one of the MMR genes and acquire inactivation of the second wild-type allele in their tumors, fulfilling the Knudson's two hit hypothesis for inactivation of tumor suppressor genes. Somatic inactivation of the corresponding wild-type allele occurs almost exclusively through point mutations or (partial) gene loss; bi-allelic inactivation then leads to complete abolition of

the protein function. This results in a defective DNA MMR system, since the MMR proteins are involved in the correction of single nucleotide mismatches and small insertions or deletions that may arise during DNA replication (Lawes DA.,2002).

The mechanism of the MMR complex has been largely elucidated and is depicted in Fig. 2.



**Figure 2.** The human Mismatch repair system (From *Clin Cancer Res.*, 2012, 18(6): 1506-1512).

The mismatch repair system was first studied in bacteria in which three proteins, MutS, MutL and MutH, were identified. In humans, at least seven mismatch repair genes are involved in mismatch repair and their names derive from their structural homology to the bacterial proteins: the MutS homologues (MSH), MSH2 on chromosome 2p16, MSH3 on chromosome 5q11, and MSH6 on chromosome 2p16; the MutL homologues (MLH), MLH1 on chromosome 3p21, and MLH3

on chromosome *2p16*; and post-meiotic segregation homologues (PMS), PMS1 and PMS2 on chromosome *7p22*. No MutH homologues have been identified in humans (*Hegan DC., 2006*). MSH2 and MSH6 bind together to form a heteroduplex (MutS $\alpha$ ) that predominantly identifies single base mispairs, while MSH2 and MSH3 (MutS $\beta$ ) combine to identify short insertions or deletions. MSH2 is essential for both complexes to function, while a functional overlap exists between MSH3 and MSH6. MLH1 and PMS2 (MutL $\alpha$ ) or MLH3 (MutL $\gamma$ ) also bind together to form a heteroduplex that interacts with MutS $\alpha$  or MutS $\beta$  complex, stimulating excision and resynthesis of the abnormal DNA. Similarly to MSH2, also MLH1 is essential for both complexes to repair mismatches. Altogether, this group of four proteins recruits exonuclease-1 (EXO1), the proliferating cell nuclear antigen (PCNA), DNA polymerase (Pol  $\delta$  or Pol  $\epsilon$ ), two replication factor (RPA and RFC), and a ligase, to repair DNA on the daughter strand at the mismatch point. If any of the four major proteins (MSH2, MLH1, MSH6, or PMS2) is functionally inactive, mismatches are not repaired (*Jun S., 2006, Jirincy J., 2006*). Consequently, a defective DNA MMR system increases the mutation rate and makes the cell vulnerable to mutations in genes controlling cell growth (tumor suppressor genes and oncogenes), resulting in an increased cancer risk.

In case of a defective MMR system, mutations occur frequently in small (usually mononucleotide or dinucleotide) repetitive DNA sequences, known as microsatellites. In MMR-deficient tumor cells the number of microsatellite repeat units can deviate from the corresponding normal DNA; the number of repeats is usually decreased even though it is occasionally found increased (*Sinicrope FA, 2012*).

Length or size microsatellite variation is known as MSI (microsatellite instability). MSI (formerly referred to as MIN, or RER, replication error) is the molecular hallmark of LS since approximately 95% of all LS-associated cancers show MSI (*val Lier MGF, 2010*). Although most microsatellite sequences are located in non-coding sequences (telomeres and centromeres), many genes contain repetitive sequences in their coding regions and some of these genes play key roles in the

regulation of cell growth (*Pedestrians M., 2001*). In fact, mutations in the TGF $\beta$ RII and TCF-4 genes, that normally inhibit cell growth, and in the IGF-RII and BAX genes involved in the apoptotic process (*Wang.Y., 1997*), particularly predispose to colon cancer. Moreover, the presence of polyadenine traits in the coding sequences of the minor mismatch repair genes, MSH6, MLH3 and MSH3, makes the same MMR genes targets of the MIN phenotype (*Loukola A., 2000; Plaschke J., 2004*). Identification of an even growing number of guide genes and target genes of the mutator phenotype can lead to discover new complex molecular mechanisms that underlie the process of colorectal tumorigenesis (*Alhopuro P, 2011*).

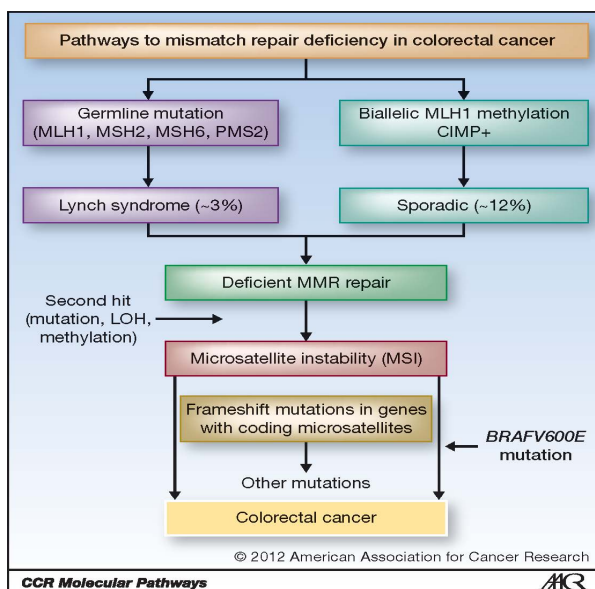
MSI thereby serves as a reliable phenotypic marker of MMR deficiency in order to pre-select patients eligible for germline mutation analysis in the MMR genes (*Zaanan A., 2011*).

However, despite the fact that MSI is a reliable marker for MMR deficiency, its specificity for LS is low since 15% of sporadic CRCs also display an MSI phenotype. This is mainly caused by somatic hypermethylation of the MLH1-gene promoter. Methylation of the MSH2 promoter has also been reported but it is to be considered as a heritable somatic methylation because it is caused by a deletion of the last exon of EPCAM that is adjacent to MSH2 on chromosome 2 (*Sinicrope FA., 2012*).

Hypermethylation of CpG islands in the MLH1 promoter (CIMP phenotype) causes severe inhibition of gene transcription thereby mimicking an inactivating gene mutation. If both copies of the gene are inactivated (bi-allelic hypermethylation), the MLH1 function is lost. This leads to microsatellite unstable cancers, especially in older patients. Therefore, in MLH1-deficient microsatellite-unstable tumors MLH1 hypermethylation can be assessed to distinguish sporadic CRCs from LS-related cancers. Moreover, recent findings have also identified the BRAF gene as a marker to distinguish LS from sporadic cases of colon cancer (*Imai K., 2008; Sharma, 2010*). It encodes a serine-threonine kinase involved in the activation of Ras/Raf/MEK signaling cascade. Specific activating mutations in the BRAF oncogene, usually the V600E missense mutation, can be detected in 40–87% of all sporadic microsatellite unstable tumors (*van Lier*

MGF., 2010). An oncogenic BRAF mutation has been described only in one case among several LS tumors (Wang. L., 2003)

These results indicate that in cases of MSI with suspicion of HNPCC BRAF mutations closely correlate with MLH1 promoter methylation in sporadic MSI CRCs, in contrast with germline mutations in the MMR genes (Fig.3).



**Figure 3.** Two molecular pathways can lead to CRC with MSI. (From *Clin Cancer Res.*, 2012, 18(6): 1506-1512).

In 1997 the National Cancer Institute recommended a panel, known as the "panel of Bethesda," comprising five microsatellites: two mononucleotide repeats (BAT25, BAT26) and three dinucleotide repeats (D2S123, D17S250, D5S346) (Boland CR., 1998). Tumors showing instability at two or more of these repeats (40% of markers) are defined as high instability (MSI-H); those with instability between 20-40% are classified as low instability (MSI-L) (Vilar E., 2014); tumors without alteration (20% or less) are classified as stable (MSS). Subsequently, in order to improve the sensitivity rate and the predictive specificity, Bethesda guidelines were revised and other loci were enclosed in the panel test: BAT-25 and BAT-26 besides three

other quasi-monomorphic mononucleotide repeats, namely NR21, NR22 and NR24 (Xicola RM., 2007, Humar A., 2004; Suraweera N., 2002).

MSI testing is also very important because several evidences suggest that MSI-H tumors (stage II) are associated with a favorable prognosis when patients are not treated with 5-fluorouracil compared to MSI-L and MSS CRC (Sargent DJ., 2010, Kim JH., 2014). These different features are probably related to the lymphocytic infiltrate characteristic of MMR-deficient tumors that determines an antitumor immune response which may be abrogated by the immunosuppressive effects of the chemotherapy (Sinicrope FA., 2012).

Besides MSI testing, analysis of MMR protein expression by immunohistochemistry (IHC) is routinely performed to identify patients with suspected Lynch syndrome. IHC testing is a specific (100%) and sensitive (92,3%) screening tool to identify MSI-H tumors (van Lier MGF., 2010; Kheirleiseid EA., 2013).

In conclusion, LS is characterized by a high lifetime risk for tumor development, especially in the case of CRC (20–70%), endometrial cancer (15–70%) and other extracolonic tumors (15%). These extracolonic malignancies include carcinomas of small intestine, stomach, pancreas and biliary tract, ovary, brain, upper urinary tract and skin. Identification of *MMR* gene mutation carriers is critical for improving cancer surveillance and effectiveness of prevention. Before *MMR* genes and their causal role in hereditary CRC cancer were identified, the *International Collaborative Group on hereditary non-polyposis colorectal cancer* had established the Amsterdam criteria I in 1990. These criteria were used to identify families eligible for molecular analysis. Subsequently modified guidelines (Amsterdam criteria II) were designed to include extracolonic LS-related cancers (Tab. 3). Nevertheless, Amsterdam criteria resulted to be very restrictive and failed to identify a large portion of *MMR* gene mutation carriers. To overcome this issue, Bethesda guidelines, which were less restrictive and had a sensitivity greater than 90% even with a lower specificity (25%), were later defined (Tab.4).

**Classical criteria (Amsterdam criteria I)**

- There should be at least three relatives with colorectal cancer One should be a first-degree relative of the other two
- At least two successive generations should be affected
- At least one colorectal cancer should be diagnosed before the age of 50 years
- Familial adenomatous polyposis should be excluded in colorectal cancer cases
- Tumours should be verified by pathological examination

**Revised criteria (Amsterdam criteria II)**

- There should be at least three relatives with an HNPCC-associated cancer (colorectal cancer, cancer of the endometrium, small bowel, ureter or renal pelvis)
- One should be a first-degree relative of the other two
- At least two successive generations should be affected
- At least one colorectal cancer should be diagnosed before the age of 50 years
- Familial adenomatous polyposis should be excluded in the colorectal cancer cases
- Tumours should be verified by pathological examination

**Table 3.** The Amsterdam Criteria

**Tumours should be tested for microsatellite instability in these situations:**

- Individuals with cancer in families meeting the Amsterdam criteria Individuals with two HNPCC-related cancers, including synchronous and metachronous colorectal cancer or associated extracolonic cancer (endometrial, ovarian, hepatobiliary or gastric cancer, or small bowel adenocarcinoma, or transitional carcinoma of the renal pelvis or ureter)
- Individuals with colorectal cancer and a first-degree relative with colorectal cancer and/or HNPCC-related cancer and/or a colorectal adenoma in which one of the cancers was diagnosed at age less than 45 years and the adenoma was diagnosed at age less than 40 years
- Individuals with colorectal cancer or endometrial cancer diagnosed at age less than 45 years
- Individuals with a right-sided colonic cancer with an undifferentiated pattern on histopathological examination diagnosed at age less than 45 years
- Individuals with signet ring cell-type cancer diagnosed at age less than 45 years
- Individuals with adenomas diagnosed at age less than 40 years

**Table 4.** The Bethesda Guidelines



### **1.3 New insights into the molecular features of Lynch Syndrome**

Recent data suggest that, in the course of evolution, in addition to the post-replicative repair, MMR proteins have developed various other functions that are highly relevant in carcinogenesis (*Jiricny J., 2006*). These new roles include:

1. DNA damage signaling caused by exogenous carcinogens (heterocyclic amines, oxidative agents and UV radiation) that is achieved through a synergistic action between the p53-homologous proteins (p53, p63 and p73) and the MutS $\alpha$ -MutL $\alpha$  complex; furthermore, in response to an exogenous damage, MLH1 interacts with the protein MRE11, a component of the "BRCA1 associated surveillance complex" (BASC), and regulates the cell cycle and the apoptotic pathway (*O 'Brein V., 2006; Yamake K., 2007*);
2. prevention of reparative recombination (gene conversion) between non-identical sequences; (*Nicholson S., 2000; Zhang J., 2006*);
3. promotion of meiotic crossover; several studies in *S. Cerevisiae* and knock-out mice have shown that homologous chromosome recombination during meiosis is controlled by MMR proteins, in order to avoid mutational events due to deletions, insertions or mismatched bases. Among the MMR proteins, MLH1, PMS2 and MLH3 are involved in this process. In fact, experimental murine deficiency of one of these three proteins is associated with male infertility (defective spermatogenesis) (*Cannavo E., 2005, Ji G., 2012*).
4. immunoglobulin diversification based on the "somatic hypermutation" (SHM) process, which is regulated by the MutS $\alpha$ -MutL $\alpha$  complex, in combination with two other proteins, AID (activation-induced cytidine deaminase) and pol  $\eta$  (DNA Polymerase "error-prone") (*Wiesendanger M., 2000; Li Z., 2006, Jiang C, 2012*); in particular, MutS $\alpha$  deficiency is associated with neoplastic transformation of T lymphocytes (*Roa S., 2010*).
5. expansion of repeated triplets (CTG, CGG) that underlie the pathogenesis of various neurodegenerative diseases such as Huntington's Disease, Myotonic Dystrophy and Fragile X Syndrome. This mechanism is still unknown, however experimental evidences indicate that, although MutS $\beta$  binds these expansions, the repair is prevented by looping conformations of these regions (*Tome S., 2009*).

Since the triplet expansion is at the basis of the anticipation of the disease in the family, loss of function of MutS $\beta$  may have a protective role against the intergenerational instability (*Dragileva E., 2009; Seriola A., 2011*).

6. modulation of microRNA biogenesis by interaction of MMR proteins with the Microprocessor complex; in particular, MutLa specifically binds to pri-miRNAs and to the complex Drosha/DGCR8 in order to stimulate the processing of pri-miRNAs to pre-miRNAs in a manner dependent on MutLa ATPase activity (*Mao G, 2012*).

These new features indicate that MMR deficiency strongly affects cellular resistance to reparative and/or apoptotic response to DNA damage because impairment of post-replicative MMR complex is associated with impairment of components of other cell systems.

### **1.4 Genotype-phenotype associations in Lynch syndrome: Canonical features**

Until now, about 500 different mutations have been identified in the MMR genes that predispose to HNPCC ([www.insight-group.org](http://www.insight-group.org)).

Germline mutations in MLH1, MSH2 and MSH6 account for approximately 40%, 39% and 11%, respectively, of all the mutations reported whereas PMS2, MLH3 and MSH3 contribution to cancer onset is less significant (*Desai TK., 2008*).

An updated list of the mutations identified has been published by the *International Collaborative Group on HNPCC* ([www.insight-group.org](http://www.insight-group.org) site), enclosing 256 mutations in MLH1, 249 in MSH2, 70 in MSH6, 13 in PMS2 and 11 in MLH3.

Mutations are distributed unevenly along each MMR gene, denoting the absence of mutational "hot spot" events. Even the nature of the germline alterations is varied.

Absence of redundant functions for MSH2 and MLH1 proteins stresses the importance of these two genes; therefore, mutations in these genes are associated with aggressive forms of HNPCC, characterized by early age of onset, typically around 45 years of age, high penetrance and high degree of microsatellite instability (MSI-H) (*Hsieh P., 2008*). The CRC incidence is similar in subjects with mutations in MLH1 and MSH2 (84% and 71% respectively);

however, individuals with alterations in the MSH2 gene show a higher incidence (48-61%) of extracolonic malignancy (endometrial, gastric, ovarian and kidney cancer) than those carrying mutations in the MLH1 gene (11-42%) (*Koornstra J., 2009*).

The clinical phenotype is different when minor genes are involved. Mutations in MSH6, for example, seem to cause a form of "attenuated" HNPCC, characterized by lower penetrance, later age of onset, usually around 60 years of age and low microsatellite instability (MSI-L) (*Lucci-Cordisco E., 2001*).

Defects in the PMS2 gene are instead associated with early tumor development and microsatellite instability, although some features are different with respect to cancers caused by MLH1 and MSH2 mutations. PMS2 mutations are associated with combined presence of multiple colorectal adenomas and glioblastomas (Turcot syndrome) (*Van Meir EG., 1998*). The specificity of brain tumor is probably linked to the accumulation of mutations in target genes (oncogenes, tumor suppressor) more specifically expressed in the brain (*Chao E., 2006*). However, the need of multiple mutational events in distinct tissues (colon, brain) would explain the low penetrance and the rarity of this syndrome.

In the MSH3 gene, missense, silent and intronic variations have been mainly identified; these mutations are associated with a severe phenotype in the case they are inherited in combination with each other, or associated with variants in the MSH2 gene (*Duraturro F., 2011*). In fact, MSH3 knockout mice showed a low susceptibility to cancer development that caused late-onset colorectal cancer, whereas double mutant MSH3-MSH6 mice showed a very similar phenotype to that found in mice lacking MSH2 (*Kuraguchi M., 2001*). These results are justified by the redundant function of the the MSH3 and MSH6 genes (*Huang J et al., 2001*). Moreover, MSH3 inactivation is primarily associated to instability of tetranucleotide repeats (EMAST) that has been frequently observed in moderately or poorly differentiated adenocarcinomas as well as in other cancers including lung, kidney, ovarian and bladder cancer (*Haugen., 2008; Lee SY., 2010*).

Similarly, mutations in the MLH3 gene are associated with a delayed

onset of disease but also with a more severe phenotype (MSI-H) in the case of co-existing MSH6 mutations (*Liu H.-X., 2003*).

In recent years, numerous studies have found an association between the development of hematopoietic and intestinal tumors in infant age and the presence of homozygous mutations in the MLH1, MSH2, MSH6 and PMS2 genes (*Bandippallian P., 2005; Herkert JC., 2011*). This phenotype was also associated with heterozygous mutations in two or more MMR genes, suggesting a mechanism of compound heterozygosity (*Plon SE., 2011; Poley JW., 2007; Peters A., 2009*).

In a subset of LS patients, a germline mutation at the 3' end of the EPCAM (TACSTD1) gene has been identified resulting in allelic-specific methylation and transcription silencing of MSH2, which is located upstream of the EPCAM gene. EPCAM gene encodes the Epithelial Cell Adhesion Molecule protein that is involved in cell signalling, migration, proliferation, and differentiation. Accordingly, this mutation may contribute to the development of extracolonic cancers (*Kang SY., 2014*).

### **Non-canonical features:**

Recently a group of Lynch-like syndrome patients was described (*Buchann DD., 2014; Lynch HT., 2014*). This group may account for as much as 70% of suspected Lynch syndrome subjects. Unlike sporadic MSI cancer, Lynch-like patients are nearly impossible to differentiate from Lynch patients: they are MSI-positive and cancer tissues express abnormal MMR protein, not only for MLH1 as in sporadic MSI cancers but also for the other MMR proteins, such as MSH2, MSH6 and PMS2, as in true Lynch syndrome cancers. Lynch-like patients show a mean age of onset comparable to LS. The only differentiating features between these two syndromes are the lower standardized incidence ratios for CRC and for non-CRC LS associated cancers in Lynch-like syndrome compared with the Lynch syndrome, and the absence of an identifiable DNA MMR gene germline mutation in Lynch-like syndrome. There are likely three potential reasons for cancer onset in Lynch-like patients: (a) a genetic process within the tumors other than germline mutations coupled with second allele inactivation, (b) unknown germline mutations in other genes than the

DNA MMR genes that can drive MSI, and/or (c) unidentified germline mutations in the DNA MMR genes (*Carethers JM., 2014; Boland R., 2013*).

*Mensenkamp AR et al. (Gastroenterology, 2014)* noted that a considerable number of MSI-positive tumors lack any known molecular mechanism for their development. Patients were screened for somatic mutations and for loss of heterozygosity in MLH1 and MSH2 genes. This research identified two somatic mutations in 13 of 25 tumors, 8 of which were MLH1-deficient and 5 were MSH2-deficient, indicating that such acquired mutations underlie more than 50% of the MMR-deficient tumors that have not been found associated with germline mutations or promoter methylation. This is in contrast with LS that is associated to germline mutations in the MMR genes.

Moreover, other hereditary factors might play a role in tumor development. For example, deletions affecting genes that regulate MSH2 degradation were shown to lead to MMR deficiency and undetectable levels of MSH2 protein (*Diouf B., 2011*). Moreover, cells lacking SETD2 (H3K36 trimethyltransferase SET domain containing protein 2) display MSI due to the loss of an epigenetic histone mark that is essential for the recruitment of the MSH2-MSH6 complex. Whether these mechanisms lead to MSH2-deficient colorectal cancer remains to be shown (*Li F., 2013*).

In these cases high-throughput sequencing procedures play an important role to identify new constitutive and somatic mutations in putative genes associated with hereditary predisposition to cancer (*Zhang J., 2011, Duraturo F., 2013*).

It is also noteworthy that, in addition to canonical inactivation via gene mutation, MMR activity can also be modulated by changes in MMR gene expression. This type of alteration may be the result of mutations occurring in regions that are not always routinely analyzed such as the promoter and the 5' and 3'-untranslated regions.

Previous studies have defined and characterized the core promoter regions of hMSH2 (from -300 to -17 upstream of the start codon) (*Iwahashi Y., 1998*) and hMLH1 (from -220 to -39 upstream of the start codon) (*Ito E., 1999*); subsequent studies have been carried out to

demonstrate that germline mutations in these regions are involved in HNPCC (*Mrkonjic M.; 2007 Raptis S., 2007*).

Regarding mutations in the 3'UTR of MMR genes, a 3-nucleotide (TTC) deletion in the MLH1 3'UTR was found in leukemia patients (*Mao G., 2008*). This alteration was shown to destroy a binding site for miR-422a and, as a result, there is a down-regulation suggesting a possible role for the miRNA in regulation of MLH1 expression (*Mao G., 2012*).

Therefore, cell levels of MMR are likely to be subject to tight regulation in order to prevent that the overproduced protein may sequester other factors involved in controlling the mutation rate. Potentially adverse consequences of overproduced MLH1 and MSH2 are highlighted by a report showing that apoptosis is induced in a human cell line when these two genes were expressed under the control of the cytomegalovirus (CMV) promoter. One possible explanation is the capture by MLH1 and MSH2 of proteins was crucial for cell cycle progression such as PCNA, a Proliferating Cell Nuclear Antigen protein, involved in DNA synthesis (*Zhang G., 1999*).

The dangerous excess of MMR protein can also be the effect of homodimerization complex as shown by a study in yeast cells of *Shcherbakova et al. (Mol. Cell. Biol., 2001)* showing that the MLH1-MLH1 homodimer replaced the MLH1-PMS1/PMS2/MLH3 heterodimer, inactivating also the MutS $\alpha$  and MutS $\beta$  functions, thus resulting in nonfunctional MMR complex.

This concept is also partially extended to other minor MMR genes: overexpression of MSH3 gene in cultured mammalian cells selectively inactivates MutS $\alpha$  because MSH2 is sequestered into a MSH2-MSH3 (MutS $\beta$ ) complex, resulting in reduced MutS $\alpha$ -dependent repair of base-base mismatches and a strong base substitution mutator phenotype (*Marra G., 1998*).

Finally, several MSI tumors with unknown cause of MMR inactivation could display a miRNA down- or up-expression genotype, that specifically modulate MMR genes (*Landau DA., 2011, Dong Y., 2014*). miRNA expression are in turn regulated by DNA damage (*Wang Y., 2013*). miRNAs able to regulate the mismatch

repair function are miR-155 and miR-21, that significantly down-regulate the core MMR proteins, MSH2, MSH6 and MLH1, and have been associated with a mutator phenotype, in particular with MSI inflammatory bowel diseases (IBD) CRCs (*Valeri N.,2010; Svrcek M.,2013*).

### **1.5 Characterization of the "variants of uncertain significance" (VUS) in the MMR genes**

Regardless of the site where a mutation occurs, the type of mutation also can make difficult a genotype-phenotype association.

Several mutations identified in the MMR genes are missense, silent or intronic variants. The influence of these variants on the development of cancer is often a controversial topic, therefore they are classified as "VUS," Variant of Uncertain Significance (*Syngal S., 1999; Couch FJ., 2008*).

Several criteria can be applied to assess the possible pathogenicity of a VUS, (*Goldgar DE, 2008; Plon SE., 2008*); these criteria are listed below: 1) de novo appearance; 2) segregation with the disease; 3) absence in normal individuals; 4) change of amino acid polarity or size; 5) occurrence of the amino acid change in a domain that is evolutionary conserved between species and/or shared between proteins belonging to the same protein family (in silico analysis); 6) effects on splicing or on protein function; 7) loss of the non-mutated allele due to a large deletion in the tumor DNA (loss-of-heterozygosity [LOH]); 8) loss of protein expression in the tumor; 9) evaluation of MSI in tumor tissue. All studies conducted to date show that none of the above criteria, including functional assays, is an indicator of pathogenicity, if taken alone; it is necessary that a combination of strategies be used in combination in order to lead to a correct assessment of the pathogenicity of uncertain variants. According to these observations, a classification of MMR sequence variants identified by genetic testing has been proposed based on a 5-class system, using a multifactorial likelihood model (Table 5).

Class	Description	Probability of being Pathogenic
5	Definitely Pathogenic	>0.99
4	Likely Pathogenic	0.95–0.99
3	Uncertain	0.05–0.949
2	Likely Not Pathogenic or of Little Clinical Significance	0.001–0.049
1	Not Pathogenic or of No Clinical Significance	<0.001

**Table 5.** Proposed classification system for MMR variant interpretation (*Colon cancer Family Registry 2009, InSiGHT Variant Interpretation Committee 2011*).

Variant-Class 5 includes coding sequence variation resulting in a stop codon (nonsense or frameshift), splicing aberration variants by mRNA assay, large genomic deletions or duplications, abrogated mRNA/protein function variants based either on laboratory assays, on evidence for co-segregation with disease and on MSI tumor and/or loss of MMR protein expression.

Variant-Class 4 includes IVS+1 or IVS+2 mutations resulting in splicing aberrations, variants abrogating mRNA/protein function based on laboratory assays, evidence of co-segregation with disease or MSI tumor and/or loss of MMR protein expression.

Variant-Class 3 includes large genomic duplications, missense alterations, small in-frame insertions/deletions, silent variants, intronic variants, promoter and regulatory region variants for which insufficient molecular evidence are available, and with intermediate clinical effects or low penetrance alleles.

Variant-Class 2 includes synonymous substitutions and intronic variants with no associated mRNA aberration, with a proficient protein expression/function, lack of co-segregation and/or MSS tumor.

Variant-Class 1 includes variants reported in control reference groups and excluded as founder pathogenic sequence variant.

According to this classification, most of the VUS tested for the MMR genes are likely to be pathogenic and thus they can be associated with the HNPCC phenotype.

For the MLH1 gene, 52 out of 73 VUS are resulted to be pathogenic (70%), similar pathogenicity has been demonstrated for 25 out of 35



VUS identified in the MSH2 gene, (71%) ([www.insight-group.org](http://www.insight-group.org)). For minor MMR genes, percentage of pathogenic VUSs is reduced due to the milder mutational contribution of these genes to the development of the disease. For the MSH6 gene, only 1 out of 8 variants studied (13%) was found to have aberrant effects on protein function; for the PMS2 gene, 4 variants were analysed and all (100%) seem to have a causative role in Lynch syndrome; for the MLH3 gene, however, functional assays have not identified any variant with certain pathogenetic significance; finally, for the MSH3 gene relevant functional studies have still not been reported ([www.insight-group.org](http://www.insight-group.org)).

### **1.6 Probability of a "synergistic effect" between low risk allelic variants in the MMR genes**

With the advent of high-throughput technologies it has been possible to analyze a great number of polymorphic variants in large cohorts of cases and controls of specific cancers, such as breast, prostate and colorectal cancer, providing new insights into common mechanisms of carcinogenesis. In some cases, VUSs make a more substantial overall contribution to cancer risk than the well-assessed severe mendelian variants. It is also possible that the simultaneous presence of some polymorphisms and VUSs in cancer predisposition genes that behave as low-risk alleles, might contribute in a cooperative manner to increase the risk of hereditary cancer (*Duraturio F., 2013*). Therefore, current literature data suggest that a significant proportion of the inherited susceptibility to relatively common human diseases may be due to the addition of the effects of a series of low frequency variants of different genes, probably acting in a dominant and independent manner, with each of them conferring a moderate but even detectable increase in the relative cancer-risk. Therefore, several functional studies based on GWAS data related to cancer susceptibility have been performed in an attempt to demonstrate the effective association and to test the hypothesis of synergistic effects between low risk allelic variants.

In a recent study on yeast genome, it has been shown that the minor alleles of the MMR complex cause a weak mutator phenotype;

however, their interaction causes a more severe mutator phenotype (Martinez SL., 2010). In this study, 11 polymorphisms and 14 missense variants of uncertain significance previously identified in the MSH2, MLH1, MSH6 and PMS2 genes, were studied by complementation tests. The mutator effect of these variants was tested singly and in combination with each other.

In 2011, Kumar *et al.* showed that some variants occurring in domain I of the MSH2 gene in yeast strains (*msh2Δ1*) behave as weak alleles in the presence of a functional protein MSH6, as they do not alter the stability of the MutSa complex. However, by combining these variants with weak alleles falling in the N-terminal region (NTR) (DNA binding domain) of the MSH6 gene, a strong mutator phenotype was found. Moreover, the mutator synergistic effect is also found between different systems of DNA damage response. A recent population study by Smith *et al.* (2011) has shown that the simultaneous presence of mutations in the TP53 gene and single nucleotide polymorphisms (SNPs) in genes belonging to different repair systems as BER, NER, MMR and DSB (Double-Strand Break Repair) complex, is associated with an earlier age of onset of breast cancer (<50 years). Therefore, in this case the authors suggest an additive or multiplicative effect.

The additive effect of low penetrance genes could also be the cause of atypical Lynch syndromes such as familial CRC type X (Lindor NM., 2005). With respect to LS, the familial CRC type X are more often located in the distal colon, extracolonic cancers are less frequent than in LS, and the age of onset is delayed. The *sine qua non* condition for this diagnosis is the absence of molecular genetic evidence of LS (MSI, IHC, or MMR mutations).

### 1.7 Scientific hypothesis and aim of the work

Molecular characterization of patients with a clinical diagnosis of Lynch Syndrome currently relies on the identification of point mutations and large rearrangements by DHPLC and MLPA, respectively, in the major MMR genes, MLH1 and MSH2.

This strategy does not always provide informative results for genetic counseling. Indeed, we analyzed 117 families selected according to international diagnostic criteria (*Amsterdam Criteria* and *Bethesda Guidelines*). Our study led to the identification of the molecular defect in 53 families while no MLH1/MSH2 mutations were identified in the remaining 64 families. Moreover, in several patients we were able to identify genetic variants of uncertain significance (VUS) (missense, intronic and silent variants).

This thesis has been focused on the characterization of the VUS identified in the major MMR genes. This has been achieved using a combination of the following approaches: segregation analysis, population studies (to exclude the polymorphic nature of the variant), assessment of Microsatellite Instability (MSI) in tumor tissues, detection of loss of protein expression in tumor tissues by immunohistochemical analysis (IHC), in silico analysis by a variety of bioinformatics tools such as HSF (Human splicing Finder), PolyPhen (Polymorphism Phenotyping), SIFT (Sorting Intolerant From Tolerant), and PredictProtein, direct analysis on the mRNA by Real-Time PCR to study gene expression modification and/or quantitative protein analysis by western-blot.

According to literature data, our study has been based on a combination of different approaches in order to verify the pathogenicity of these uncertain variants.

Analysis of two mutations detected in the 3'untranslated regions (3'UTR) of MLH1 and MSH2 genes have led to very interesting results. On the basis of in silico analysis, we had hypothesized that mutations in this regions may impair binding of putative transcriptional factors or microRNA involved in regulation of gene expression. By gene expression studies we have been able to demonstrate the pathogenetic significance of these variants.

In this study we have also analyzed the minor MMR genes, MSH6,

PMS2, MLH3, and MSH3, for germline variants detected in patients negative for germline mutations in the major MMR genes. Recently, literature data indicated that MMR proteins have other functions besides the post-replicative repair, that could be highly relevant in carcinogenesis (cell cycle and apoptosis regulations, *Jiricny J. 2006, Ji G., 2012, Mao G., 2012*) and could involve the minor MMR genes, as well. Therefore, we speculated that genetic variants in these genes could have an important role in carcinogenesis progression. Several VUS were also identified in minor MMR genes. and phenotypic characterization was carried out for some of these variants.

Finally, since many of the subjects analyzed showed co-inheritance of different genetic alterations in the MMR genes, we assume a likely additive role of low penetrance alleles in the disease development, in favor of a putative polygenic inheritance for Lynch syndrome, in according to recent literature data (*Martinez SL., 2010*).

In conclusion, the aim of this study was to clarify the pathogenetic significance of many genetic variants identified in MMR genes. This will allow to clarify the genotype-phenotype correlations in these patients, in order to improve the genetic counselling and, consequently, the clinical surveillance.

## **2. MATERIALS AND METHODS**

### **2.1 Patients**

Sixty-four families of Italian origin of which 40 families classified according to the Amsterdam Criteria and 24 atypical Lynch families selected according to MSI (Bethesda Guidelines), without well-defined pathogenetic germ-line point mutations or large rearrangements in the major MMR genes, MLH1, MSH2, were recruited from several Clinical Centers in Campania (Southern Italy). Samples from all families participating to the study were collected after informed consent of the participants.

### **2.2 DNA, RNA and protein isolation**

Total genomic DNA was extracted from 4-mL peripheral blood lymphocytes collected with EDTA using a BACC2 Nucleon Kit (Amersham Life Science), according to the manufacturer's recommendations, and from formalin fixed and paraffin embedded (FFPE) tumor/normal tissues by standard method. For FFPE extraction, tissue sections of 25 µm were previously cut with a microtome and then used. Each section (usually 4 for specimen) was dewaxed in 1 mL of xylene, then dipped in 1 mL of lysis solution containing 0.1 M Tris-HCl, 0.01 M EDTA, 1 M NaCl, 1% SDS, 400 µg / mL of Proteinase K overnight at 48°C. DNA was extracted with a phenol/chloroform/isoamyl alcohol solution (25: 24: 1, v/v), subsequently with a chloroform/isoamyl alcohol solution (24: 1, v/v), and, finally, DNA was precipitated with two volumes of absolute ethanol and 1/3 volume of 3M sodium acetate, pH 5.5. After incubation overnight at -20°C, the pellet was washed with ethanol at 70% and resuspended in sterilized TE buffer (Tris 10mM pH7.5-EDTA 1mM pH8).

Total genomic RNA was extracted from 4 mL peripheral blood lymphocytes collected with EDTA with TRIZOL solution (Qiagen) according to the manufacturer's protocols.

After extraction, both DNA and RNA were quantified by spectrophotometer analysis and their integrity was verified by electrophoretic analysis on 1% agarose gel.

Total protein extracts were obtained from the same blood sample used for RNA isolation according to the Qiagen protocol.

Evaluation of protein concentration was performed by spectrophotometer analysis ( $\lambda = 595\text{nm}$ ), according to the Bradford method using the Bio-Rad Protein Assay Reagent (Bio-Rad Laboratories).

### 2.3 Microsatellite Instability (MSI) analysis

The analysis of microsatellite instability was performed on DNA extracted from tumor and normal tissue/peripheral blood lymphocytes, with the GeneQuality CC-MSI kit (AB-Analytical).

Through the use of oligonucleotides labeled with different fluorophores (HEX, FAM and TAMRA), 4 different amplified products with a single capillary electrophoresis running was performed, allowing simultaneous analysis of 12 microsatellite markers, since each mix allows amplification of 3 microsatellites, as shown in Table 6.

Microsatellite	Multiplex	Localization	Sequence	Length	Fluorophore
TGF $\beta$ RII	2	3p22	A(n)	80-90	FAM
NR21	1	14q11.2	T(21)	101-109	TAMRA
BAT40	2	1p12-13.3	A(40)	120-128	HEX
NR24	1	2q11.2	T(24)	132-145	FAM
D18S58	1	18q23	CA(n)	144-160	HEX
D2S123	2	2p15	CA(n)	206-220	FAM
D5S346	4	5q22-23	CA(n)	102-124	FAM
BAT25	3	4q12	T(25)	120-132	HEX
D17S250	3	17q12	CA(n)	144-158	FAM
BAT26	4	2p21-22	A(26)	202-225	FAM
TH01	4	2p23-2pter	TCAT(n)	154-178	HEX
TPOX	3	11p15-15.5	TGAA(n)	232-248	HEX

**Table 6.** Microsatellites analysis with the GeneQuality CC-MSI kit (AB-Analytical).

The panel of amplified markers comprises five quasimonomorphic mononucleotide repeats, BAT-25, BAT-26, D2S123, D5S346, D17S250 (Bethesda Panel), four mononucleotide repeats, NR21, NR24, BAT40, and TGF $\beta$ RII, one dinucleotide marker D18S58 and two tetranucleotide repeats, TPOX and TH01, with the latter two markers used as internal controls and not for evaluation of microsatellite instability.

For each amplification, the reaction mixture is composed as follows:

Buffer 10X (2.5  $\mu$ L); MgCl<sub>2</sub> (0.75  $\mu$ L); dNTPs (0.5  $\mu$ L); primer mix (1  $\mu$ L); SuperABTaq (0.2  $\mu$ L); H<sub>2</sub>O (up to volume); DNA (20-25 ng) in a final volume of 25  $\mu$ L. The amplification program is the following: 95°C for 8 min, 10 cycles at 94°C (30 sec), 60°C → 55°C (45 sec with 0.5°C/cycle decrement) and 72°C (30 sec), 22 cycles at 94°C (30 sec), 55°C (45 sec) and 72°C (30 sec), and a final extension at 72°C (10 min).

The electrophoretic separation by capillary electrophoresis on an ABI 3130 Genetic Analyser (Applied Biosystems) is based on the molecular length of each fragment and on the fluorophore type, as shown in Table 10. Before electrophoresis, 1  $\mu$ L of each amplicon was added to 24  $\mu$ L of deionized formamide (Life Technologies) and 1  $\mu$ L of size standard (GeneScan 500 ROX, Life Technologies), and incubated for 5 minutes at 95°C. After the electrophoretic run, normal and tumor multiplex-PCR electropherograms for each patient were compared by the GeneScan software (Applied Biosystems).

## **2.4 Immunohistochemistry analysis**

Immunohistochemistry analysis was performed on an Benchmark XT automatized immunostainer (Ventana Medical Biosystems, Tucson, USA). The antibodies tested were:

- 1) Anti-MLH1, mouse monoclonal clone M1 (Ventana);
- 2) Anti-MSH2, mouse monoclonal clone G219-1129 (Ventana);
- 3) Anti-MSH6, mouse monoclonal clone 44 (Ventana).

The detection system used was a Ventana DAB-iView which is based on the Streptavidin- Biotin-conjugated revelation system.

3-5 microns FFPE tumoral tissue sections were cut and let stick on electrostatically charged glass slides (SuperFrost) overnight at 42°C.

Glass slides were loaded on an immunostainer Ventana BenchMark XT for automatic staining. Firstly the instrument operates a dewaxing and subsequently an incubation with the inhibitor of endogenous peroxidase (H<sub>2</sub>O<sub>2</sub>). The subsequent steps of antigen unmasking (HIER, Heat Induced Epitope Retrieval) and the antibody (Ab) incubation were setted as follows:

	MLH1	MSH2	MSH6
HIER	90 min. at 99°C pH 8	60 min. at 99°C pH 8	90 min. at 99°C pH 8
Ab incubation	40 min. at room temperature	32 min. at 37°C	32 min. at 37°C

Analysis was performed with the detection system “iView DAB” which consists firstly in dispensing the biotinylated secondary antibody and then the enzyme-conjugated streptavidin (HRP, horseradish peroxidase). The antigen positive complexes are detected by the addition of the DAB chromogen (diamminobenzodine) and its substrate ( $H_2O_2$ ). The samples are finally counterstained with the nuclear dye hematoxylin. After immunostaining, slides are taken by the instrument, dehydrated, added with xylene and few drops of conditioner and mounted with a cover glass.

Nuclear staining was observed with an optical microscope with positivity represented by the presence of brown staining. This positivity was compared to blue nuclear epitopes, in which the specific antigen was not present. The internal positive control was represented by lymphocytes, stroma and functional mucosal crypts, while the negative control was obtained by slides without primary antibody.

## **2.5 MMR germline point mutations and large genomic rearrangements analysis**

For MMR (MLH1 GeneBank *NG\_007109.2*, MSH2 GeneBank *NG\_007110.2*, MSH6 GeneBank *NG\_007111.1*, PMS2 GeneBank, *NG\_008466.1*, MLH3 GeneBank *NG\_008649.1*, and MSH3 GeneBank *NG\_016607.1*) germline point mutations analysis, DNA amplification reactions were performed with oligonucleotide primers chosen in order to include all exons and intron-exon boundaries of each gene. All oligonucleotides were obtained with PrimerBlast software ([www.ncbi.nlm.nih.gov/tools/primer-blast/](http://www.ncbi.nlm.nih.gov/tools/primer-blast/)). PCR reactions were performed in a total volume of 50  $\mu$ L containing 5  $\mu$ L of 10X PCR buffer (Roche), 200  $\mu$ M of each dNTP, 25 pM of each primer, 1.5 mM of  $MgCl_2$ , 2 U of FastStart Taq DNA polymerase (Roche) and 100 ng of genomic DNA. PCR conditions were as follows: initial denaturation at 95°C for 4 min, followed by 35 cycles with denaturation at 95°C for 30 sec, annealing for 30 sec at the melting temperature of each primer, extension at 72°C for 45 sec, followed by a final extension step at 72°C for 7 min. The PCR products were analyzed on a 1-2% agarose gel in Tris-acetic acid (TAE)-EDTA 1X standard buffer and visualized by ethidium bromide staining.

Identification of sequence variants within each amplicon was performed by DHPLC (Denaturing High Performance Liquid



Chromatography) analysis, using the Wave 3500HT system (Transgenomic), according to the manufacturer's recommendations. Before dHPLC analysis, the PCR products were denatured at 95°C for 5 min and gradually cooled to 20°C using a temperature ramp of 1°C/min on a PCR thermocycler to enable efficient formation of heteroduplex. The mobile phase gradient and running column temperature selected for optimal heteroduplex separation were determined for each amplicon using the Wave Marker 4.4 software provided with the instrument. In cases of samples showing abnormal patterns on DHPLC, DNA sequencing in both the forward and reverse directions was performed using the ABI 3100-Avant automatic DNA sequencer (Applied Biosystems, CA). For nucleotide numbering, the first A of the initiator ATG codon is nucleotide +1 of the MMR mRNA sequence.

Detection of large genomic deletions/duplications was performed by MLPA (Multiplex Ligation-dependent Probe Amplification) analysis. MLPA is a semi-quantitative multiplex PCR, performed using the SALSA MLPA P003-B1 MLH1/MSH2 kit (MRC-Holland) according to the manufacturer's instructions. Fragment analysis was conducted on an ABI Prism 3130 Genetic Analyser using the GeneMapper software (Applied Biosystems). Migration of fragments was calculated by comparison to the GeneScan LIZ-500 size standard (Applied Biosystems). Peak areas were then exported to a Microsoft spreadsheet

([www.mrc-holland.nd](http://www.mrc-holland.nd).) and calculations were done according to the method described by Taylor et al. (*Human Mutation*, 2003). A 30–50% decrease in the peak area(s) indicated a deletion of the corresponding exon(s), while a 30–50% increase in the peak area(s) indicated a duplication of the corresponding exon(s). MLPA results were confirmed in at least two independent experiments and by cDNA sequencing.

## **2.6 RNA based analysis**

For qualitative RNA analysis, MMR cDNA was synthesized using 1 µg of total RNA, 500 µg of Oligo(dT)<sub>12-18</sub>, and 1 µL of Superscript III reverse transcriptase (Invitrogen), in the presence of 4 µL 5X RT buffer, 1 µL DDT (0.1 M) and 1mM dNTPs. The reaction was run for 50 min at 42°C in a 20 µL of reaction volume, heated to 70°C for 15 min and snap-chilled on ice.

1  $\mu$ L of each cDNA was amplified with primers that produce several overlapping fragments covering all exons. The PCR reaction was performed as described above for DNA PCR.

PCR products were analyzed on a 8%/10% polyacrylamide gel in Tris-Boric acid-EDTA (TBE) 1X standard buffer and visualized by silver staining. Bands of different lengths, corresponding to different splicing isoforms, were cut and eluted from gel in 30  $\mu$ L of sterile water according to the procedure of the QIAquick Gel Extraction kit (Qiagen). Purified bands were amplified for sequencing analysis.

Quantitative RNA analysis was performed by Real Time PCR on an CFX96 Real-Time System (Bio-Rad-Laboratories). PCR products were verified on agarose gel before performing the relative quantification. A calibration curve to assess the efficiency of the PCR reaction was performed on at least three serial dilutions (1:10) of the reverse transcriptase products. Each Real Time PCR (100-150 bp long) was performed in triplicate in a 20  $\mu$ L reaction mix containing 12.5  $\mu$ L of 2X *SYBR Green I* PCR Master mix (Bio-Rad Laboratories), 0.38  $\mu$ L of a 20  $\mu$ M primer mix, 2  $\mu$ L of cDNA (5ng/ $\mu$ L) and 7.12  $\mu$ L of nuclease-free water. The cycling conditions consisted of an initial denaturation step at 95°C for 3 min, followed by 40 cycles (95°C for 15 sec, 62°C for 30 sec, 82 °C for 20 sec) and 80 cycles performed according to standard protocols for melting curve analysis. CT values were determined by automated threshold analysis and data were analyzed with the CFX Manager software version 2.1 (Bio-Rad Laboratories). The relative expression of the target transcript was calculated with the comparative Ct method using a cDNA fragment from the glucoronidase (GUS) housekeeping gene as control. Three several forward and reverse primers for *MLH1* cDNA quantification were carried out, by amplifying fragments spanning between exons 3-5, 4-5 and 13-14. Two forward and reverse primers for *MSH2* cDNA quantification were designed, by amplifying fragments spanning between exons 4-5 and 13-14 (Tab.7).

## **2.7 Protein analysis**

Western blot analysis was performed on 30-50  $\mu$ g of total protein extracts. Proteins were denatured at 95°C for 10 min in a buffer containing 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v)  $\beta$ -mercaptoethanol, 75 mM Tris/HCl pH 6.8, 0.001% (w/v)

Fragment name	Primer sequences	Size amplification
3 FP cMLH1	CCAGTATTCTACCTATGGCTTTCGACGTG	
5I RP cMLH1	GGTTTAGGAGGGGCTTTCAG	198bp
4 FP cMLH1	CAACGAAAACAGCTGATGGAA	
5II RP cMLH1	GATCTGGGTCCCTTGATTGC	103bp
13 FP cMLH1	GCAGGGACATGAGGTTCTCC	
14 RP cMLH1	GCTTGGTGGTGTGAGAAGG	169bp
4 FP cMSH2	ACCGTTGTTGAAAGGCAAA	
5 RP cMSH2	TTGATTACCGCAGACAGTGATG	100bp
13 FP cMSH2	TGGTGACAGTCAATTGAAAGGA	
14 RP cMSH2	CCCATGCTAACCCAAATCCA	152bp
GUSFP	GAAAATATGTGGGTTGGAGAGCTCATT	
GUSRP	CCGAGTGAAGATCCCTTTTA	120 bp

**Table 7.** Primer sequences and amplicon size for MLH1 and MSH2 mRNA quantification.

bromophenol blue and separated by SDS-page electrophoresis in a 10% Acrylamide/Bis 29:1(v/v) gel (Tris/HCl 1.5 M pH 8.8, SDS 10%) for 2 hours at 100 volts. Proteins were then transferred on a nitrocellulose membrane (GE Healthcare) in a Tris-Glycine-SDS (Tris 25 mM, Glicine 192 mM, SDS 0.1% ) buffer.

Proteins were visualized by reversible staining with a solution of 0.1% Red Ponceau S (Sigma-Aldrich) in 5% acetic acid solution.

After elettroblotting, non specific sites on the nitrocellulose membrane were saturated over night at 4°C with a milk solution (5% w/v) containing 0,1% Tween 20 (w/v) in TBS 1% (Tris HCl 1mM pH 7.5, NaCl 15 mM). Filters were incubated over-night at 4°C with the following antibodies: mouse monoclonal anti-MSH2, clone GB12, (Calbiochem) diluted 1:100, and mouse monoclonal anti-MLH1, cloneG168-15 (BD Pharmingen) diluted 1:250 and subsequently normalized with a mouse antibody anti  $\alpha$ -actin (dil 1:1000). Membranes were washed with 1x TBS-Tween 20 buffer for 5 minutes (3 times) and incubated for 45 minutes with a secondary antibody (anti-mouse dil 1:10000) conjugated to peroxidase (Sigma). The antigen-antibody complexes were visualized by the ECL-Immobilon chemiluminescence reagents (Millipore) and subsequent autoradiography. Western blots bands were quantified by the ImageJ software.

## 2.8 In silico analysis

In silico analysis was performed for all MMR gene variants of uncertain significance, not yet reported in literature, using the

following bioinformatics tools:

- Human Splicing Finder (HSF) ([www.umd.be/HSF/](http://www.umd.be/HSF/)) for missense, silent and intron variants;
- Phenotyping Polymorphism (PolyPhen)
- (<http://genetics.bwh.harvard.edu/pph/>) for missense variants;
- Sorting Intolerant From Tolerant (SIFT)
- (<http://blocks.fhcrc.org/sift/SIFT.html>) for missense variants;
- Predict Protein (<https://www.predictprotein.org/>) for missense variants.

HSF software is a new tool used to predict the effects of mutations on splicing signals or to identify splicing motifs in human sequences. It contains all available matrices for auxiliary sequence prediction and also presents a new position weight matrix to assess the strength of 50 and 30 splice sites and branch points.

PolyPhen, SIFT, and PredictProtein algorithms were used to perform structural analysis of missense point mutations in order to evaluate the functional activity of the mutated protein. Predictions are based on a combination of phylogenetic, structural and sequence annotation information with the substitution position in the protein. The range-scores calculated for each program are as follows:

<b>PolyPhen</b>	Benigne $\Delta \leq 0.5$	Probably damaging $0.5 < \Delta \leq 1.5$	Possibly damaging $\geq 1.5$
<b>SIFT</b>	Not affected $>0.05$	Affected $<0.05$	—
<b>PredictProtein</b>	Neutral effect $<-50$	$<-50$ score $<50$ weak effect	$>50$ strong effect

In silico analysis was also performed to identify regulating factors of gene expression and their target motifs, in gene sequences such as promoter and untranslated regions. The following bioinformatics tools were used:

- TRANSFAC(<http://www.biobaseinternational.com/product/transcription-factor-binding-sites>) to search DNA sequences for putative transcription factor binding sites;
- TargetScan 6.2 (<http://www.targetscan.org>), miRanda (<http://www.microrna.org>), miRDB (<http://mirdb.org>) that predict biological targets of miRNAs by searching for the

presence of conserved 8mer and 7mer sites that match the seed region of each miRNA.

## 2.9 Functional analysis

The MSH2 3'UTR variant (c.\*226 A>G) identified at the heterozygous state in a HNPCC family was tested by the Luciferase Reporter assay.

The MSH2 3'UTR of a carrier of this mutation was amplified by PCR with a primer pair containing XhoI and NotI restriction sites, respectively. Oligonucleotide sequences were as follows:

XhoI-3'UTR MSH2 <i>forward</i>	ATACTCGAGAAAATCCCAGTAATGGAATG
NotI-3'UTR MSH2 <i>reverse</i>	ATAGCGGCCGCTCAAATTCCACAACTACA

The adapter bases are indicated in red, while the restriction sites are indicated in blue. The whole amplicon sequence (296bp) was tested for restriction enzyme sites by the GeneJockey II software.

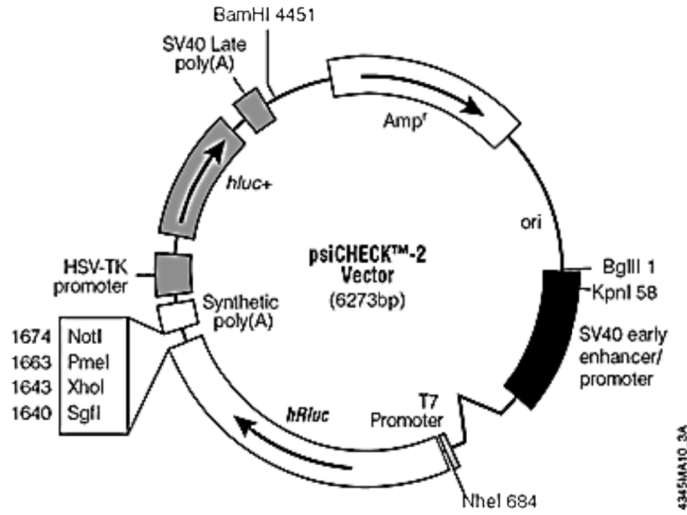
PCR reaction was performed as above described (*paragraph 2.5*) at a melting temperature of 60°C.

The PCR product was cloned into the PSICHECK2 vector (*Promega*) (Fig. 4), downstream from the Renilla luciferase coding region (hRluc).

The wild-type and the mutated constructs were transfected in SW480 cells (MMR+ human colon adenocarcinoma), using Lipofectamine 2000 (Invitrogen) as transfectant agent. These cell lines grow as adherent cells in RPMI 1640 (Gibco) supplemented with 10% fetal bovine serum (Sigma), 1% GlutaMAX-I (Invitrogen) and 1% Penicillin/Streptomycin (Gibco). The cultures were maintained at 37°C in a 5% CO<sub>2</sub>-humidified atmosphere and they were kept sub-confluent (70-80%) for the subsequent transfection experiments.

24 hours before transfection adherent cells were plated into 96-well plates at a density of 25 x 10<sup>3</sup> (70-90% confluence) in 100 µL of Optimem Medium (Invitrogen) in the absence of serum/antibiotic.

Subsequently, two solutions were prepared, containing 30ng/µL of plasmid diluted in Opti-MEM medium (24µL) and 0,5 µL of Lipofectamine diluted in Opti-MEM medium (24,5 µL), respectively.



**Figure 4.** Graphic representation of the PSICHECK2 vector.

These two solutions were mixed in a single tube, and incubated 20 minutes at room temperature. This mixture was added to each well (50 $\mu$ L). The plate was incubated at 37°C. Subsequently, 50/100  $\mu$ L of OptiMEM was added to each well. The plates were incubated for six hours in the appropriate growth conditions (37 ° C, 5% CO<sub>2</sub>) and then 10% FBS was added in each well. Luciferase activity was measured at 48 h after transfection, using a dual luciferase reporter assay (Promega) according to manufacturer's instructions and performed on a 20<sup>n</sup>/20<sup>n</sup> Luminometer (Turner BioSystems). Relative luciferase activity was calculated by normalizing the *Renilla* luminescence to the firefly luminescence. The assay was performed in 5X replication in three independent experiments.

Then, HT29 (MMR+ human colon adenocarcinoma) and SW480 cells were transfected with the pre-miR miRNA precursor hsa-miR-137 (Ambion, Life Technologies) and pre-miR miRNA negative control FAM (Ambion, Life Technologies) at a final concentration of 30 nM, using Lipofectamine 2000 (Invitrogen) according to manufacturer's instructions. For qRT-PCR analysis, total RNA was extracted 48h after transfection, with TRIZOL solution (Qiagen) according to the manufacturer's protocols.

### **3. RESULTS**

In this study, we analyzed the DNA samples of 64 subjects, belonging to families diagnosed with Lynch syndrome, in which no germline point mutations or large rearrangements in MMR genes, hMLH1 and hMSH2, had been found. Variants of uncertain significance (VUS) were identified in these major MMR genes.

The 64 families were selected on the basis of the Amsterdam criteria and Bethesda guidelines, as follows:

- 40 families fulfilling the Amsterdam Criteria;
- 24 families identified according to the Bethesda Guidelines and classified as "HNPCC", on the basis of the presence of high/low microsatellite instability (MSI-H/L) and the absence of the BRAF mutation.

In these families several VUS were identified in the major MMR genes and extensive phenotypic characterization was carried out on some of them in order to investigate their potential pathogenic role; furthermore, samples negatives for pathogenic variants in major MMR genes were then analyzed for mutations in the minor MMR genes, MSH6, PMS2, MLH3 and MSH3.

#### **3.1 Phenotypic and functional characterization of VUS in the major MMR genes**

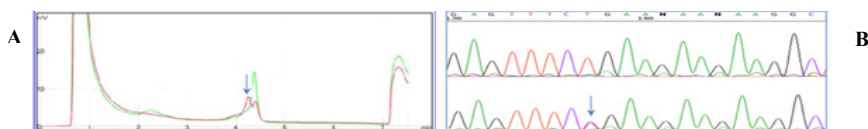
In the MLH1 gene, we performed a phenotypic and a functional characterization of 3 of the most interesting variants of uncertain significance that had been identified and associated with severe disease phenotypes (Table 8).

The missense mutation in exon 3, c.304G>A (p.Glu102Lys) (Fig.5), was found in a subject, belonging to an Amsterdam Criteria family, that had developed a tubule-villous adenocarcinoma of the colon at the age of 44 years (MSI-H).

## Results

Id patients	Exon of Gene	Nucleotide change	Aminoacid change	Reportage	MSI Phenotype	Clinical Phenotype
00-13	3 of hMLH1	c.304 G>A	p.[Glu101Valfs*14, Glu102Lys]	Ellison et al. 2001	MSI-H	Cancer of small intestine diagnosed at age 43; mother presented a colon polyp at age 73, MUT+.
05-DS	5 of hMLH1	c.438 A>G	p=Gln146	This study <sup>a</sup>	MSI-H	Right-colon cancer at age 44 and kidney cancer at age 48; sister died of colon cancer at age 53.
09-08	3'UTR of hMLH1	c.*30_32 delTTC		Viel et al. 1997	MSI-H	Right-colon cancer at age 45; father with kidney cancer diagnosed at age 60, MUT+.
<sup>a</sup> Not present in databases InSiGHT <sup>13</sup> and MMR gene variants <sup>14</sup> . Id, Identification; MUT+, mutation positive; MUT-, mutation negative. Nomenclature in according to HGVS, <a href="http://www.hgvs.org/mutnomen">www.hgvs.org/mutnomen</a> .						

**Table 8.** VUS identified in the MLH1 gene in patients with MSI-H.



**Figure 5.** Identification of the c.304G>A mutation in the MLH1 gene. (A) DHPLC analysis and (B) DNA sequencing analysis of exon 3.

This variant had been previously reported as a class-3 variant of uncertain significance. In our case, it was not possible to carry out a segregation study for this variant, due to the limited availability of family members.

Nevertheless, we were able to use alternative approaches to ascertain its pathogenic significance, according to the multifactorial likelihood model (Goldgar DE., 2008; Plon SE., 2008).

First, this variant was not identified in 52 healthy controls, thus allowing us to exclude it as a polymorphic site. Secondly, according to the fact that this variant falls in an evolutionarily conserved region, functional prediction by PolyPhen, SIFT and PredictProtein algorithms indicated that this sequence variant may affect the protein function (Fig. 6). Finally, the HSF analysis revealed the loss of a binding site for SF2/ASF (enhancer splicing factor) (Fig. 6, 7), thus reinforcing that this alteration could be responsible for a defect in the normal splicing process.

To verify this hypothesis, we extracted RNA from lymphocytes of the patient and of three normal controls. Using primers that amplify a cDNA region of MLH1 spanning exons 2-4 (forward primer: 5'-



## Results

TGTTAAAGAGGGAGGCCTGA-3') (reverse primer: 5'-TCCATCAGCTGTTTTTCGTTG-3'), we obtained a PCR product of 216 bp, corresponding to the normal allele in all samples.

Mutation	PolyPhen	SIFT	Predict Protein	HSF
<b>c.304G&gt;A</b> <b>p.Glu102Lys</b>	Probably damaged (0,999)	Damaging (0)	Affected function protein (96)	-SF2/ASF (c.302_308) +5'ss <sup>1</sup> ΔCV(c.302_306) = +67.88

**Figure 6.** In silico analysis of the c.304G>A mutation identified in the MLH1 gene.  
<sup>1</sup>A new cryptic splice site (5ss') created by the mutation (ΔCV: coefficient of variation).  
 The scores reported for each analysis were according to the score-range calculated by the algorithms (see section 2.7 of *Material and Methods*).

In addition to the fragment of the expected molecular weight, in the patient's sample we found a weaker fragment of slightly lower molecular weight, probably corresponding to an abnormally spliced allele (Fig.7).

Sequence analysis of this abnormal fragment, after elution from the gel and re-amplification, revealed the presence of a 5 bp deletion, due to the activation of a cryptic splice site (Fig. 8). Therefore, the missense p.Glu102Lys variant acts as a truncating variant since it generates a down-stream premature stop codon at codon 120 and, consequently, a loss of function at the protein level.

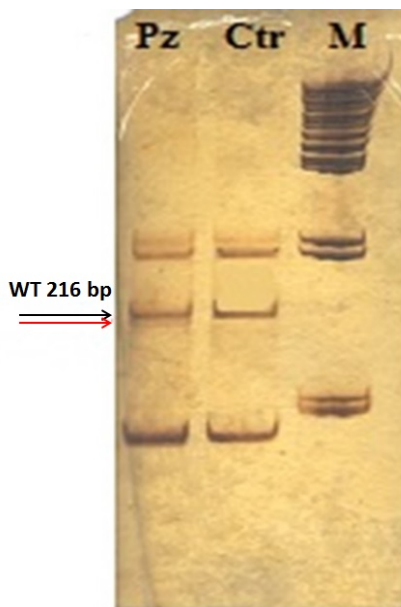
We also analyzed another MLH1VUS variant, the c.438A>G in exon 5 (Fig.9).

It is a silent variant (p=Gln146) identified in a patient who had developed colon cancer when about 50-year aged, and who belonged to a family in which several cases of abdominal tumors had been observed. Unfortunately, it was impossible to analyze other members of the family, but we ruled out the presence of the alteration in the DNA of 60 healthy subjects. Furthermore, even for this variant we performed an in silico analysis to predict its pathogenicity. The computational analysis performed using the HSF program indicated

## Results

that this nucleotide change resulted in the creation of a new binding site for SF2/ASF that could adversely affects the splicing mechanism (Fig. 10).

To verify this hypothesis, we extracted RNA from lymphocytes of the patient and of three normal controls. We amplified a region encompassing the entire MLH1 transcript using a forward primer (5'-ACGTTTCCTTGGCTCTTCTG-3') and a reverse primer (5'-AATCAATCCACTGTGTATAAACCAA-3') and in all samples we obtained a PCR product of the expected normal size (2,464 bp). In addition to this main band, we also observed the presence of several fragments of lower molecular weight (Fig. 11). These fragments were present in all samples, although in the patient's sample these abnormal

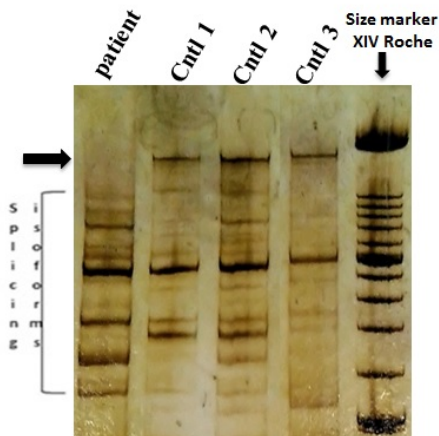


**Figure 7.** Analysis of MLH1 cDNA in a patient carrier of the c.304G>A variant. 10% polyacrilamide gel electrophoresis of the MLH1 cDNA ex2-4 fragment of the patient (pz) and of a normal control (ctr); the arrows show the normal fragment (up) and a weak abnormal fragment of slightly lower molecular weight (down); M: size marker IX (Roche).



## Results

fragments were much more represented to the full-length mRNA. Sequence analysis of these fragments allowed us to characterize all of them as products of alternative splicing of the MLH1 gene; we observed mRNA isoforms with the in-frame skipping of exons 17-18, exons 6/9/10, exon 15, exon 9/10 and exons 9/10/11, respectively.



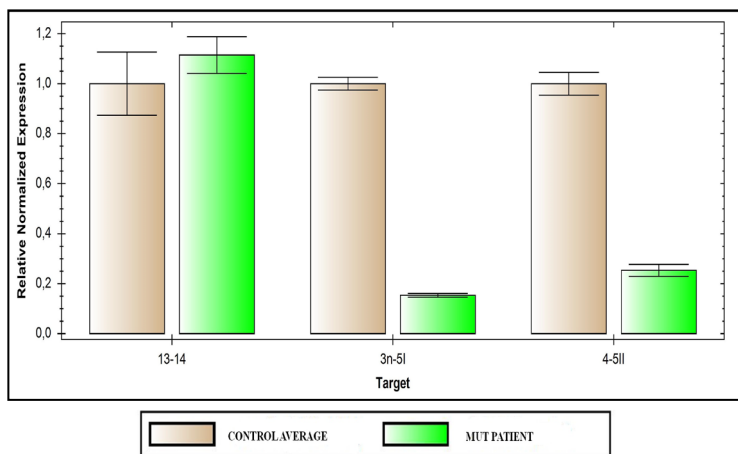
**Figure 11.** Polyacrylamide electrophoresis analysis of the MLH1 cDNA in a patient carrier of the variant c.438 A>G and in three normal controls samples; the arrow shows the amplified fragment corresponding to the full-length cMLH1; the bands below correspond to alternative splicing isoforms. Marker XIV (Roche).

To confirm this data, we set up a quantitative real-time PCR reaction in which fragments corresponding to exons 3-5, 4-5 and 13-14 of the MLH1 transcript were amplified, using GUS as a reference gene. Fragments encompassing the region in which the variant falls (3-5 and 4-5 exons fragments) were under-represented in the patient with respect to normal controls whereas expression of the 13-14 fragment showed no significant differences between patient and controls (Fig. 12). These data indicate that the creation of a new binding site for the enhancer protein SF2/ASF could promote the process of alternative splicing, thus causing the formation of a greater number of aberrant isoforms.

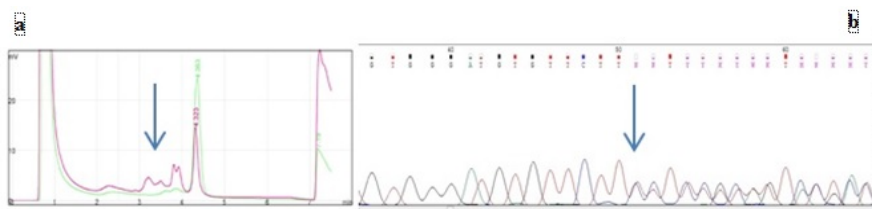
The other MLH1 variant analyzed was the c.\*30\_32delTTC (3'UTR), identified in several HNPCC families and not present in 54 healthy subjects (Fig.13). This variant was described in literature as a class-1 variant of uncertain significance. However, in our study, this alteration

## Results

was always associated with a severe phenotype (MSI-H) when it was co-inherited with another variant, the c.454-51t>c, described in literature as class-1 (table 9).



**Figure 12.** q-Real-Time PCR analysis of MLH1 mRNA in the patient with the variant c.438A>G. Gray columns represent means  $\pm$  SD of 3 different negative controls. The results represent the average of three independent experiments.



**Figure 13.** Identification of the c.\*30\_32delTTC mutation in the MLH1 gene. (A) DHPLC analysis and (B) DNA sequencing analysis of the MLH1 3'UTR.

The computational analysis performed by the HSF program for this intronic variant showed a possible deleterious effect on the splicing process (Fig.14). Therefore, we may assume for these two genetic alterations a pathogenic additive effect.

We could not apply the several bioinformatic analysis to the variant c.\*30\_32delTTC since it occurs in an untranslated region, therefore we used the TRANSFAC algorithm for prediction of TF binding sites. The results showed that the variant destroys the binding site for one of

## Results

two TF identified by the program (fig. 15). Recently, a computational analysis has identified a putative binding site for miR-422a in the MLH1 3'UTR that is disrupted by the 3-nucleotide deletion. A reciprocal feedback regulation mechanism between MLH1 and miR-422a was also demonstrated (Mao G., 2012).

	c.*30_32delTTC	c.454-51t>c	MSI Phenotype	Clinical Phenotype
Family n.1	Mut+	Mut+	MSI-H	Left and sigma colon cancer diagnosed at age 39 et 44 (relapse); squamous epithelioma (40) and sebaceous epithelioma (47); several colon cancer in early age in the family.
Family n.2	Mut+	Mut+	MSI-H	Tight-colon adenoma at age 48; father with kidney cancer at age 60.
Family n.3	Mut+	Mut+	MSI-H	Colon cancer at age 23 and relapses; several colon and extracolonic cancers in the family.
Family n.4	Mut+	Mut+	MSI-H	No-Hodjkins Lymphoma at age 50 and colon cancer at age 60; several colon cancer in the family both in early and late age.
Family n.5	Mut+	Mut+	MSI-H	Breast cancer at age 45; sister with colon cancer at age 41; mather with right colon cancer at age 74.
Family n.6	Mut+	Mut-		No cancer at age 39; several colon and extracolonic cancers at early and late age in the mather family

**Table 9.** Families with the c.\*30\_32delTTC and c.454-51t>c mutations in the MLH1 gene. Families (n.1 to n.5) with double mutation, c.\*30\_32delTTC and c.454-51t>c; family (n.6) with only the c.\*30\_32delTTC mutation.

Potential Branch Points ↑

Sequence Position	cDNA Position	Branch Point motif	CV for reference sequence (0-100)	CV for mutant sequence (0-100)	Variation
27	c.454-55	gttttAt	55.44	67.27	New site

Fas-ESS hexamers

Sequence Position	cDNA Position	Reference sequence	Set	Mutant sequence	Set	Variation
-56	c.454-56	ggtttt	2, 3			Site broken

Predicted PESS Octamers from Zhang & Chasin (New Dataset)

Sequence Position	cDNA Position	Reference motif	Motif value (0-100) reference sequence	Mutant motif	Motif value (0-100) mutant sequence	Variation
27	c.454-55	gttttatt	31.51	gtttcatt	31.24	-0.84 %
28	c.454-54	ttttatttt	79.05	tttcatttt	39.62	-49.88 %
29	c.454-53	tttattttt	75.65			Site broken -100

IIEs from Zhang et al.

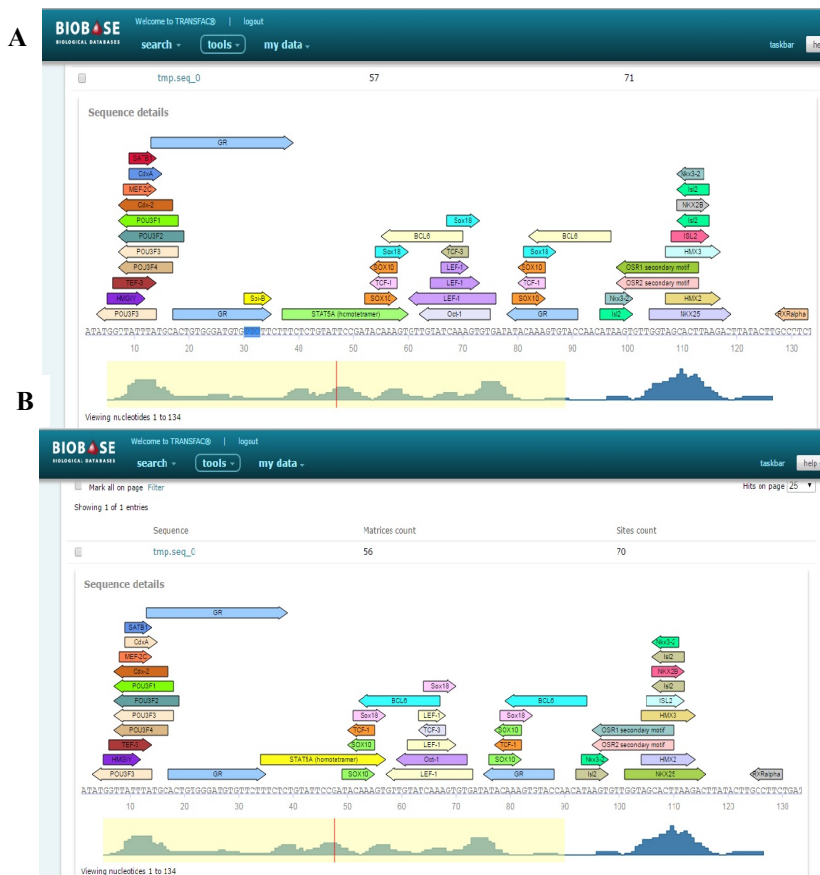
Sequence Position	cDNA Position	Silencer motif reference sequence	Silencer motif mutant sequence	Variation
26	c.454-56	ggtttt	ggtttc	
27	c.454-55	gtttta		Site broken

**Figure 14.** HSF analysis of the c.454-51t>c variant. ESS motifs by Fas-ESS algorithm: Exonic Splicing Silencer; PESS motifs by PESS-octamers algorithm: Putative Exonic Splicing Silencers; IIEs motifs: Intron-Identity Elements.

On the basis of these data, we examined the MLH1 expression at mRNA and protein levels in a patient carrier for this variant. Indeed,

## Results

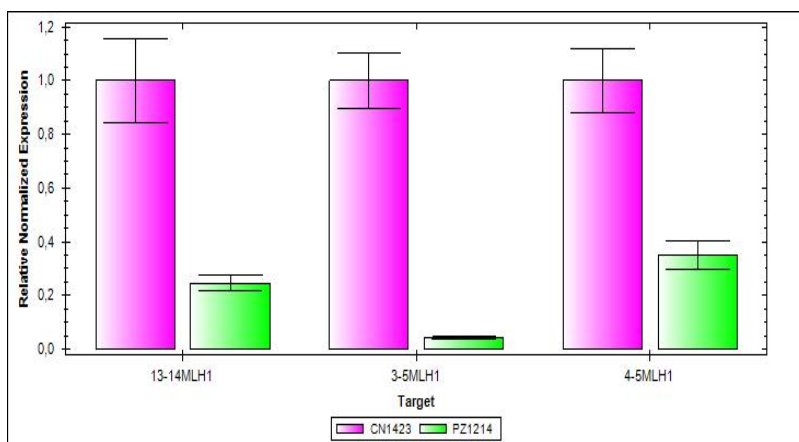
quantitative real-time PCR showed a reduction in MLH1 mRNA levels in one of the patients with the mutation (Fig. 16).



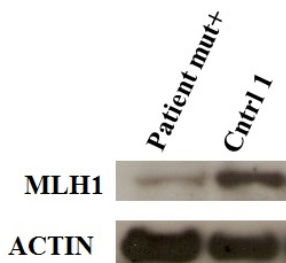
**Figure 15.** TRANSFAC analysis of the MLH1 3'UTR. (A) wild-type sequence and interaction between the region c.\*30\_32 and two factors, GR (glucocorticoid receptor) and SPIB (Tf acting in B-cell receptor signaling); (B) mutated sequence and interaction between the c.30\_32delTTC region and SPIB.

Western blot analysis of proteins isolated from the lymphocytes of patient carrier of this mutation revealed a lower level of MLH1 in the patient compared to a negative control (Fig. 17).

## Results



**Figure 16.** MLH1cDNA q- Real-Time PCR analysis in a patient with the two variants, c.\*30\_32delTTC and c.454-51t>c. Green columns correspond to results from the patient. Pink columns represent means  $\pm$  SD of 3 different negative controls. The results were the average of three independent experiments.

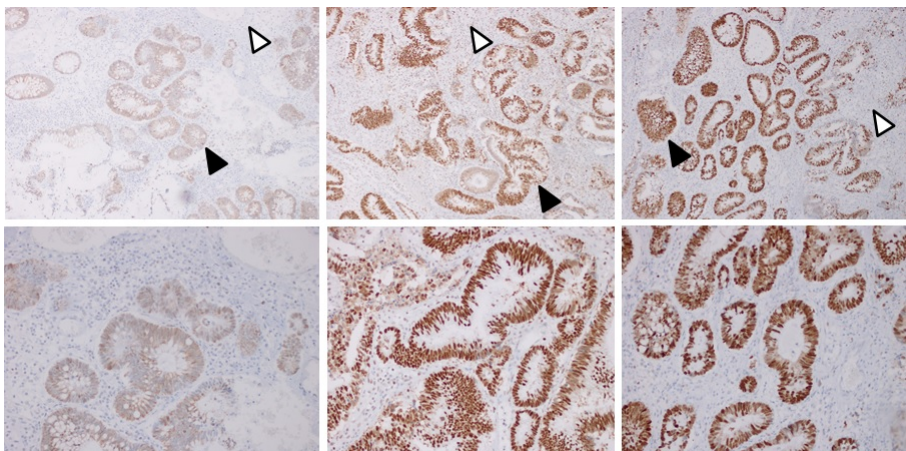


**Figure 17.** Western blot analysis of MLH1 expression in patient with two variants c.\*30\_32delTTC and c.454-51t>c. Actin was used as an internal (positive) control.

These results were confirmed by immunohistochemical analysis (IHC) on tumoral tissue sections of a patient carrier of this mutation. We observed low intensity of the nuclear staining for assessment of MLH1-protein expression compared to MSH2 and MSH6 protein expression (Fig. 18).

These results were in accordance with previous literature data (*Mao et al. 2008*) and are indicative of pathogenicity for this variant.





**Figure 18.** MLH1, MSH2, MSH6 immunohistochemistry (IHC) in a patient carrier of the c.\*30\_32delTTC and c.454-51t>c mutations. Low intensity MLH1 (1)10X and (2) 20X; strong intensity MSH2(3) 10X and (4) 20X; strong intensity MSH6 (5) 10X and (6) 20X. Colon tumour cells (filled arrow head point) and IHC<sup>+</sup> lymphocyte internal control (open arrow head point ).

We performed a multivariate analysis for 1 out of 3 variants of uncertain significance detected in the MSH2 3'UTR, the variant c.\*226A>G, that was found associated with severe phenotypes (Table 10).

This mutation was identified in three subjects of an HNPCC family and was not present in 54 healthy controls (Fig.19). It had been described in literature as class-2 of uncertain significance.

In our study, this alteration was associated with low microsatellite instability (MSI-L) and with diversified phenotypes, as showed in the pedigree of the mutation carrier (Fig.20).

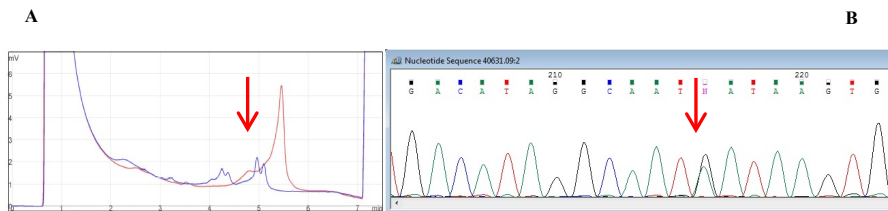
A computational analysis was performed by the TRANSFAC program for prediction of TF binding sites, and a combination of three different software algorithms (TargetScan 6.2, miRanda, miRDB ) that predict biological targets of miRNAs. Results are shown in figure 21 and 22, respectively.

According to in silico analysis, we performed a quantitative real-time PCR analysis that showed increased levels of MSH2 mRNA expression in all the three patients with the mutation (Fig. 23).

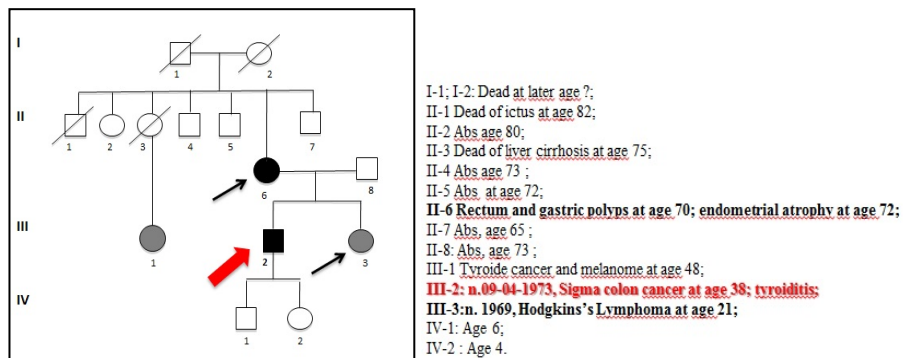
## Results

Id patients	Exon of Gene	Nucleotide change	Aminoacid change	Reportage	MSI Phenotype	Clinical Phenotype
0106	5 of hMSH2	c.917C>T	p.Ala306Val	This study	MSI-H	Rectum colon cancer diagnosed at age 32; several colon cancers in the family.
10-04	5 of hMSH2	c.942+3_16del13A	-	This study	MSI-H	Colon cancers in the family in according with Amsterdam criteria.
1335	3'UTR of hMSH2	c.+226A>G	-	Thompson et al., 2013	MSI-L	Sigma colon cancer diagnosed at age 38; sister with Hodgkin's lymphoma at age 21; mother with colon and gastric polyps at age 70 and endometrial polyp at age 72. Other colon cancers in the family.

**Table 10.** VUS identified in the MSH2 gene in patients with MSI-H and MSI-L.



**Figure 19.** (A) DHPLC analysis and (B) DNA sequencing analysis of the MSH2 3'UTR: identification of the c.\*226A>G mutation.

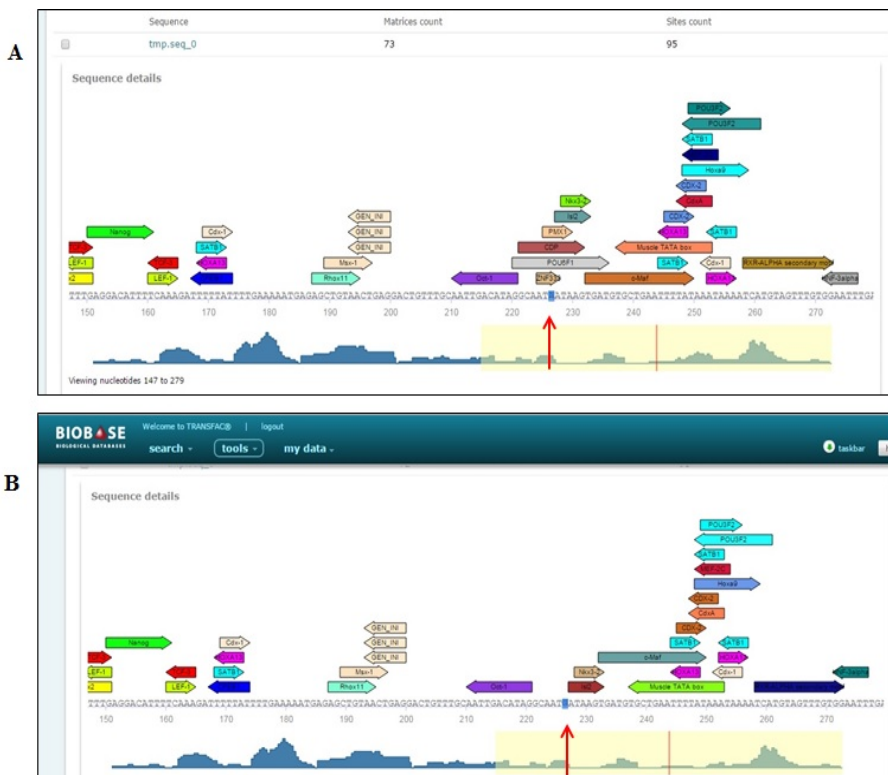


**Figure 20.** Pedigree with the segregation of the c.\*226A>G mutation. The red arrow indicates the family index case.

Western blot analysis of proteins isolated from lymphocytes of one (III-2 index-case) of the three patients showed an increased level of MSH2 in the patient compared to negative controls, in accordance with the real-time data (Fig. 24).

## Results

An immunohistochemical analysis (IHC) of FFPE tumoral tissue sections of two patients (II-6 endometrial tissue and III-2 colon tissue) revealed an increased level of MSH2 protein compared to normal tissues (Fig. 25, Fig. 26).



**Figure 21.** TRANSFAC analysis of the MSH2 3'UTR. (A) Wild-type sequence and interaction between the mutated region c.\*226A>G and four protein factors, ZNF333 (zinc-finger protein 333), POU6F1 (Tf acting in B-cell receptor signaling), CDP (cut-like homeobox 1 protein), and PMX1 (paired related homeobox 1 protein). (B) Mutated sequence and loss of interaction between the mutated region c.\*226A>G and each of the four factors. Red arrows indicate the WT (A) and Mut (B) nucleotide.

In order to determine whether the variant c.\*226A>G in the MSH2 3'-UTR altered the expression of upstream coding sequences, the WT and MUT 3'-UTR were cloned downstream the Renilla luciferase reporter gene (Fig. 27).

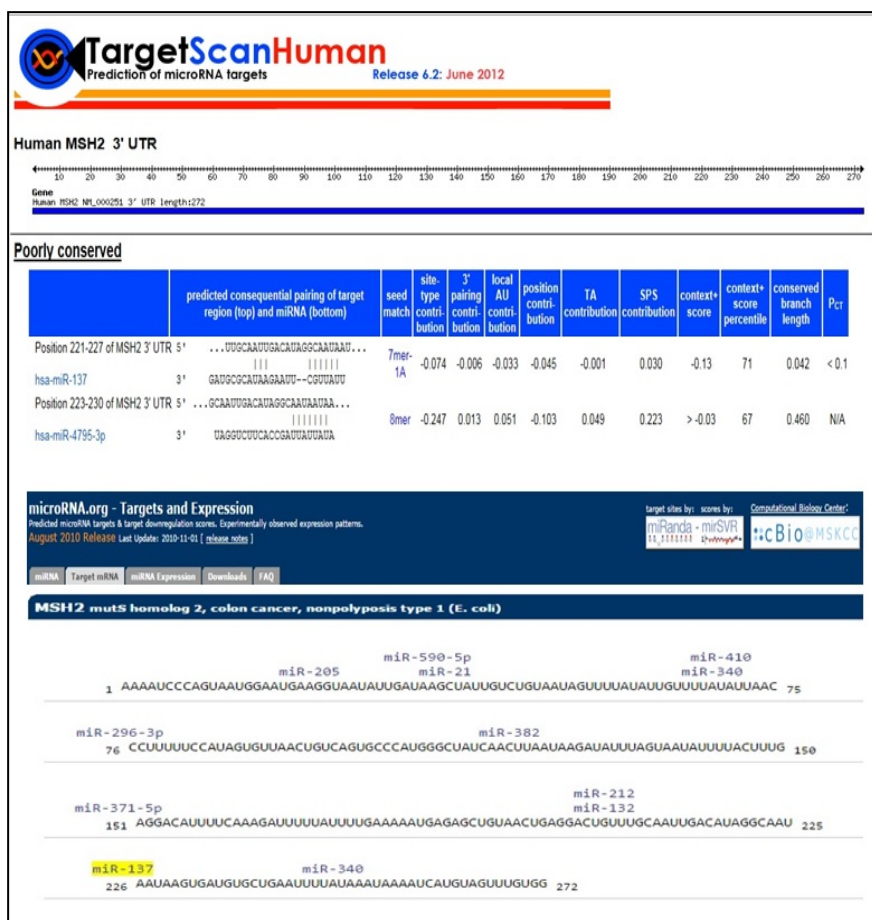
## Results

---

The reporter gene constructs were transfected into SW480 cells and 48 hours later cells were collected for luciferase assay and quantitative mRNA analysis. Results showed that the construct bearing the MUT 3'-UTR consistently induced higher luciferase activity than the construct with the WT 3'UTR (Fig. 28).

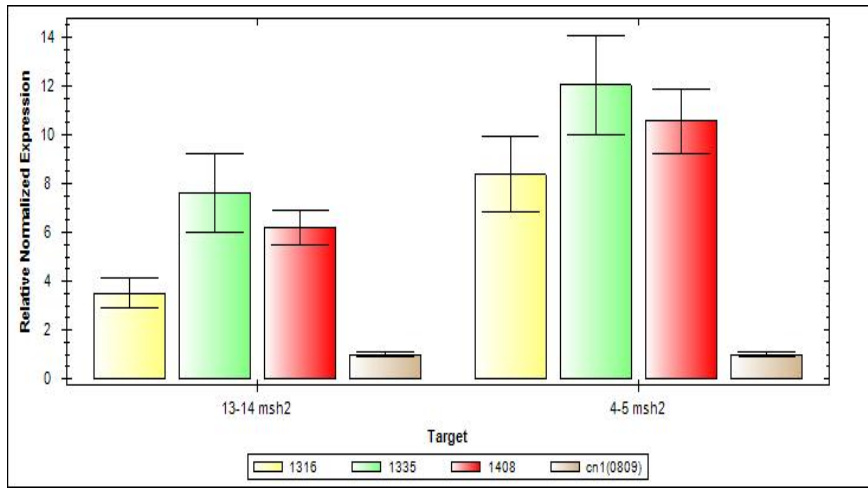
We next performed a functional study to determine whether in vitro miR-137 interacts with the MSH2 3'UTR, according to previous in silico analysis and to literature data that indicate this miRNA as a tumor suppressor in the colon, inhibiting cell proliferation. The effect of miR-137 on the MSH2 transcript was investigated in two colorectal cancer cell lines, HT29 and SW480, that express high and low levels of the miR-137, respectively (*Balaguer F., 2010*).

All transfections were performed in triplicate in 12-well plates. As shown in Fig. 29, the ectopic miR-137 expression after 48h of transfection significantly down-regulated the MSH2 mRNA levels in both cell lines particularly in HT29 cells, suggesting a possible role for this miRNA in MSH2 expression regulation.

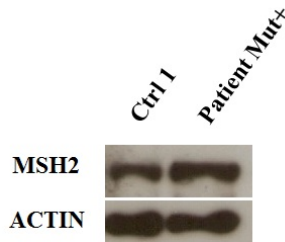


**Figure 22.** In silico analysis of the MSH2 3'UTR by the prediction tools TargetScan (up) and miRanda (down). The miRNA identified were hsa-miR-137 and has-miR-4795-3p by TargetScan, and hsa-miR-137 by miRanda.

## Results



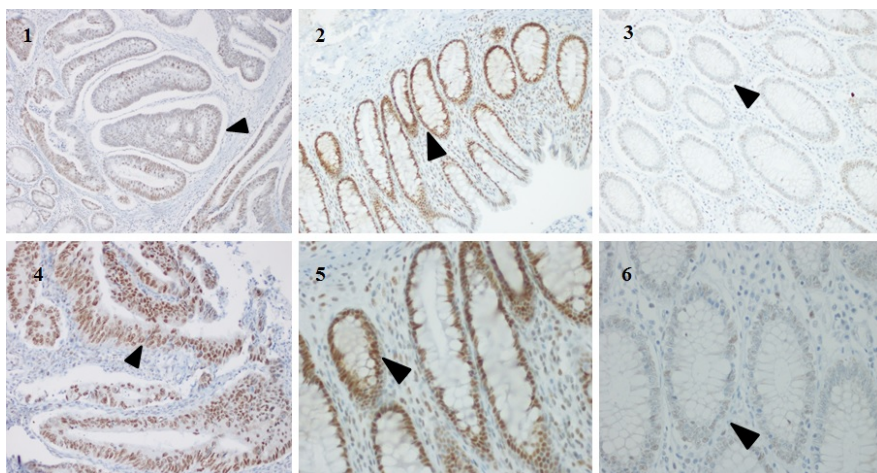
**Figure 23.** q-Real-Time PCR analysis of the MSH2 mRNA in the patients with the c.\*226A>G variant. Yellow, green and red columns represent the mRNA of II-6, III-2 and III-4 pedigree cases, respectively. Beige columns represent means  $\pm$  SD of 3 different negative controls. The results represent the average of three independent experiments.



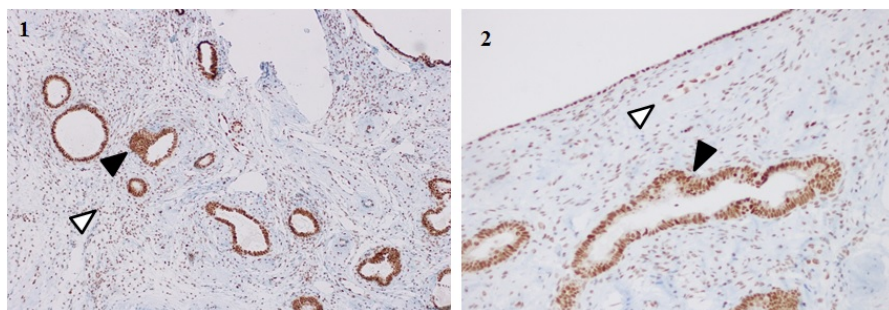
**Figure 24.** Western blot analysis of MSH2 expression in the index-patient with the c.\*226A>G variant. Actin was used as an internal positive control.



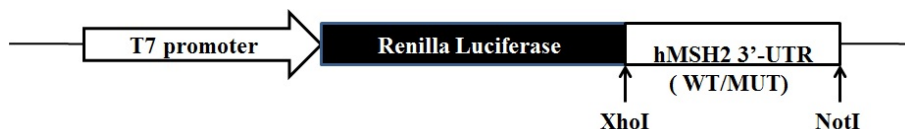
## Results



**Figure 25.** MSH2 immunohistochemistry (IHC) results in III-2 index patient. (1) Strong positive IHC (80% staining intensity) in the colon tumor cells (filled arrow head point) 10X and (4) 20X; (2) strong positive IHC (80% staining intensity) in the normal mucosa cells (filled arrow head point) of the patient 10X and (5) 20X; (3) weak positive IHC (30% staining intensity) in normal mucosa colon cells (filled arrow head point) of a control subject 10X and (6) 20X.

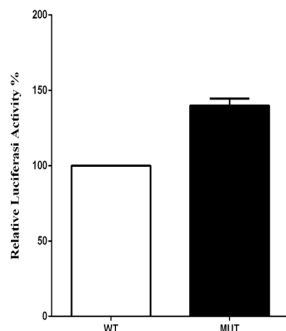


**Figure 26.** MSH2 immunohistochemistry (IHC) results in the II-6 patient. (1) Strong positive staining (100%) in the endometrial atrophic polyp cells (filled arrow head point) 10X and (2) 20X, compared with IHC+ internal stromal cells (80% staining intensity) (open arrow head point).

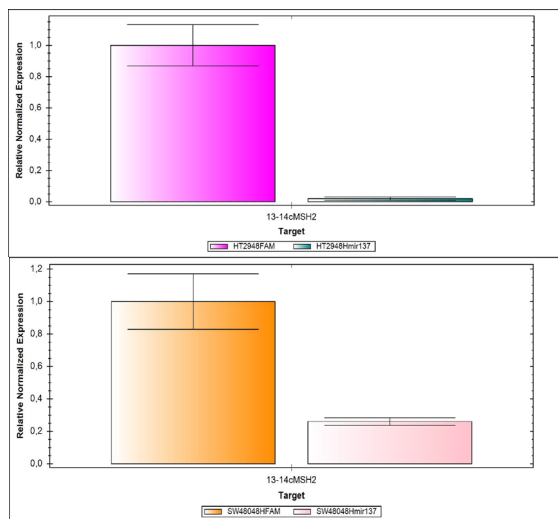


**Figure 27.** Schematic diagram of the WT and Mut luciferase reporter gene constructs.

## Results



**Figure 28.** Relative Luciferase expression. The Renilla luciferase reporter activity is significantly increased in MUT-3'UTR vector compared to WT-3'UTR in SW480 cells ( $p < 0.005$ ). The data were normalized to the Firefly luciferase activity. Values are expressed in percentage as the mean  $\pm$ SD of three replicate experiments.



**Figure 29.** q-Real-Time PCR analysis of endogenous MSH2 mRNA in the HT29 and SW480 cells transfected with hsa-miR-137 and a negative control (FAM). Results represent the average of three independent experiments.

### 3.2 Mutational analysis of minor MMR genes

Screening for point mutations was performed by DHPLC and sequencing analysis on amplicons corresponding to the MSH6, PMS2,



## Results

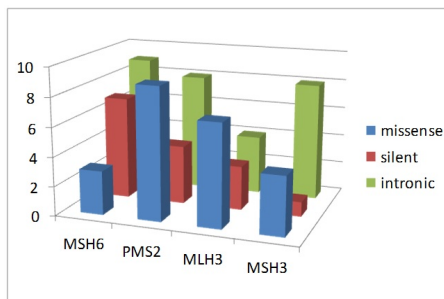
MLH3 and MSH3 coding regions while large genomic deletions/duplications analysis was carried out by Multiplex Ligation-dependent Probe Amplification (MLPA) as described in Materials and Methods.

A total of 77 variants were identified in the minor MMR genes, 5 of which were pathogenic mutations, 5 were missense variants already described in the literature as pathogenic mutations (Fig. 30) and 67 resulted to be unclassified variants (missense, silent and intronic variants) (Fig.31).

A multivariate analysis was performed for these 67 VUS, by phenotypic characterization (segregation study, analysis on 52 negative control samples, MSI and IHC analysis) and functional studies (*in silico* analysis, mRNA and protein analysis).

5 PATHOGENIC MUTATIONS:	5 MISSENSE VARIANTS ALREADY DESCRIBED IN THE LITERATURE AS PATHOGENIC MUTATIONS:
1) <b>MSH6</b> exon 4, c.2049_50 insAGT(p.insSer683) (this study)	1) <b>MSH6</b> exon 2, c.431 G>T (p.Ser143Ile) ( <a href="#">Wu et al. 1999</a> )
2) <b>MSH6</b> exon 5, c.3296_97 delTT (p.Ile1098fs) (this study)	2) <b>MSH6</b> exon 4 c.663 A>C (p.Glu221Asp) ( <a href="#">Devlin et al. 2008</a> )
3) <b>MSH6</b> exon5, c.3261_62 insC (p.Phe1088fs) ( <a href="#">Bonk et al., 2003</a> )	3) <b>MSH6</b> exon 4, c.2398 G>C (p.Val800Leu) ( <a href="#">Kolodner et al. 1999</a> )
4) <b>PMS2</b> exon 3-4dup (this study)	4) <b>MSH6</b> exon 4, c.2633 T>C (p.Val878Gly) ( <a href="#">Wiinen et al. 1999</a> )
5) <b>MLH3</b> exon7, c.184_190delGGGAGTG (p.Gly61fs) (this study)	5) <b>MSH6</b> exon 5, c.3226 C>T (p.Arg1076Cys) ( <a href="#">Plaschke et al. 2000</a> )

**Figure 30.** Pathogenic mutations identified in minor MMR genes.



**Figure 31.** Histogram of several VUS identified in the minor MMR genes.

This approach allowed us to demonstrate the likely pathogenicity of 10 out of 27 variants identified in the MSH6 gene, 3 out of 21 variants identified in the PMS2 gene, 2 out of 15 variants identified in the MLH3 gene, and 2 out of 13 variants identified in the MSH3 gene (data not shown).

The most relevant result of this mutational analysis on minor MMR genes was the detection of multiple co-inherited molecular alterations in different genes in a group of patients. Table 11 reports the subjects who resulted to be carrier of this condition, in relation to their clinical phenotype.

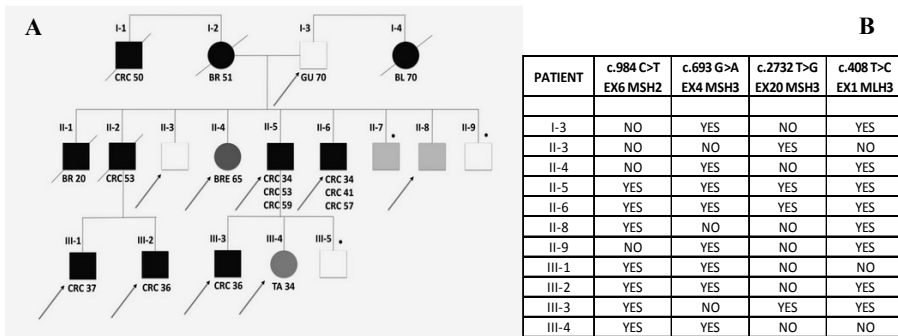
Anyway, to evaluate the pathogenic role of these variants, it would be necessary to analyze the members of each family to verify the segregation of each of these mutations with the disease. Unfortunately, this study was possible only for one family of the 64 subjects analyzed, because of the poor cooperation of the other families. The pedigree of this family is shown in Figure 32A. Mutational analysis of the MLH1 and MSH2 genes in the index case allowed us to identify the silent mutation c.984 C> T in exon 6 of the MSH2 gene. Analysis of the minor MMR genes identified the following mutations: the c.693 G> A in exon 4 and the c.2732 T> G in exon 20 of the MSH3 gene, and the c.408 T> C in exon 1 of the MLH3 gene. This analysis was extended to other nine family members. The results are reported in Figure 32B. Analysis of microsatellite instability was carried out on DNA extracted from tumor tissue and normal mucosa of subjects II-5 (proband) and III-3 (proband's affected child). This analysis showed a high degree of microsatellite instability (MSI-H) and this is in agreement with the strong mutator phenotype observed in the two patients (Fig. 33).

## Results

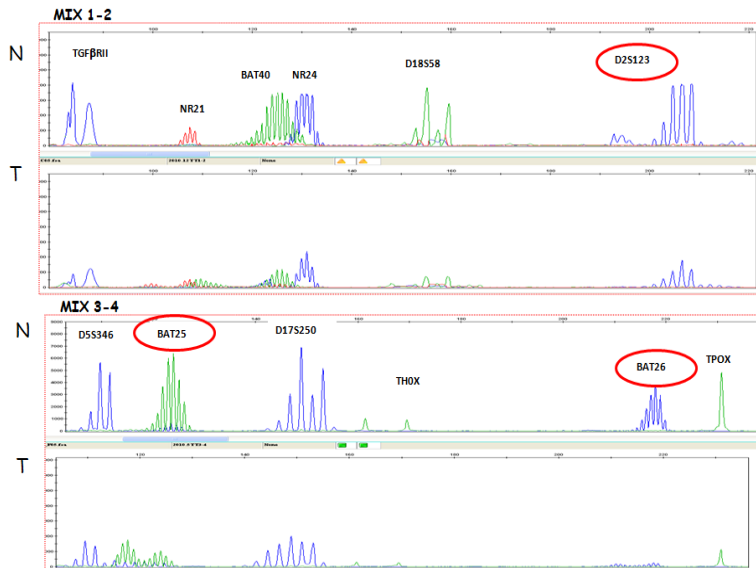
**Table 11.** Patients with more mutations in the MMR genes.

	VARIANTS	CLINICAL HISTORY	FAMILY HISTORY
06/1	<b>PMS2:</b> ex2 c.59 G>A (Arg20Gln), ex12c.2007-7c>t, ex15 c.2570 G>C (Gly856Ala), <b>MSH3:</b> ex23 c.3131 G>a(Ala1044Thr) +/-	Moderately differentiated mucinous adenocarcinoma at age 51; endometrial cancer at age 55.	Mother dead of intestinal obstruction at age 52; father dead of colon cancer at age 48; two brothers dead of colon cancer at age 42 and 50, respectively.
06/2	<b>MLH3:</b> ex4 c.3484 A>T(Ser1161Cys), <b>MSH3:</b> ex20 c.2732 T>G (Leu910Thr)	Moderately differentiated endometrial carcinoma at age 63; sigma-rectum mucinous adenocarcinoma at age 68.	Mother dead of intestinal obstruction at age 35; sister with uterus cancer at age 57; sister dead of intestinal obstruction at age 64; cases of leukemia, lymphoma and meningioma among children.
06/3	<b>MSH2:</b> ex7c.1077-10T>C <b>PMS2:</b> ex7c.780 C>G(Ser) +/-, <b>MLH3:</b> ex11 c.4263G>a (p-Gln) +/-, ex4 c.3484 A>T(Ser1161Cys)	Left colon cancer at age 70 with liver metastases.	Parents dead of old age; brother with bladder cancer at age 50; brother with lung cancer at age 72; sister with fibromatous uterus at age 49 and rectum polyp at age 59.
06/4	<b>MLH3:</b> c.184_190delGGGAGTG (Gly61fs), ex4 c.3484 A>T(Ser1161Cys) +/-, ex11c.4263G>a (p-Gln) +/-	Colon cancer at age 49; lung metastases at age 57.	Grandfather dead of intestinal obstruction at age 63; father dead of colon cancer at age 58; sister with kidney and uterus cancer at age 46; sister with sigma colon and uterus cancers at age 53; daughter with polyp at age 30.
01/3	<b>MSH6:</b> ex5 c.3261_62insC (Phe1087fs) <b>MSH3:</b> ex12 c.1860G>A (Asp>Asn), ex12 c.1763+71c>a	Right poorly differentiated adenocarcinoma at age 61.	Mother with colon-rectum cancer at age 50; brother with colon-rectum cancer at age 58; sister with colon adenoma at age 63 but negative at molecular testing.
01/5	<b>MSH6:</b> ex5 c.3295_97delTT (Ile1098fs), <b>MLH3:</b> ex1-3 c.666 G>A (Lys), ex1-9 c.2221 G>T(Val740Phe)(2003), ex1-11 c.2533A>G (Ser844Gly), ex3c.3465+37_40del tcct, ex10c.4170+13c>g(2003)	Moderately differentiated adenocarcinoma at age 46; brother not affected and negative at molecular testing; sister with colon polyp at age 30 but negative at molecular testing.	Grandmother with colon cancer at age 78; father with colon cancer at age 50; uncle with colon cancer at age 70; cousin with colon cancer at age 46; uncle with colon cancer at age 64 and his child with several cases of polyps and uterus cancer between 35 and 50 years.
04/15	<b>MSH6:</b> ex4A c.642C>T (Tyr) (1999) <b>PMS2:</b> ex13 c.2253T>C (Phe)(1998), <b>MLH3:</b> ex3 c.3465+37delT (2003), <b>MSH3:</b> ex22 c.3000-22t>a, ex22 c.3000-33_36deltgaa	Colon cancer at age 46.	Grandmother with colon cancer at age 78; father with colon cancer at age 71; uncle with colon cancer at age 64; several cases of polyps at early age among cousins.
02/10	<b>MSH6:</b> ex4S c.2941 A>G(Ile980Val), ex4D c.1164C>T (His)(1999), ex4F-G c.1395A>T (Ala)	Colon cancer at age 57.	Father with colon cancer at age 70; sister with several polyps at age 55-60.
02/11	<b>MLH1:</b> ex16 c.1856 G>A (-Lys). <b>MSH6:</b> ex4S c.2941 A>G(Ile980Val)	Sigma colon cancer at age 42; two colon adenoma at age 44.	
04/10	<b>MLH1:</b> ex19 c.2152 C>T (His718Tyr) (1997) <b>MSH6:</b> ex4Q c.2633T>C (Val877Ala) (1999), ex4C c.990A>T(Ser) <b>MLH3:</b> ex11 c.4263G>a (p-Gln)+/-	Moderately differentiated adenocarcinoma at age 62.	Mother dead of leukemia; father dead of colon cancer; three brothers with colon polyps between 61 and 71 years; sister with colon polyp at age 56; cousin dead of colon cancer at age 61; son with two colon polyps at age 30.
05/3	<b>MSH2:</b> ex1 c.211+9C>G; <b>c.1661+12G&gt;A</b> <b>PMS2:</b> ex2 c.59 G>A (Arg20Gln)(1995), ex7 c.780 C>G(Ser)+/(1995), ex12 c.2007-7c>t(2006), ex14 c.2324 A>G (Asn774Ser)(1998), ex14 c.2340C>T (Pro)(2006) <b>MLH3:</b> ex1-3 c.666 G>A (Lys)	Sigma moderately differentiated adenocarcinoma at age 53; descending colon adenocarcinoma at age 54.	Mother with colon cancer and uterus cancer at age 47; Uncle dead of sarcoma at age 42.
05/4	<b>MSH2:</b> ex6 c.984C>T (p=Ala), <b>PMS2:</b> ex7 c.780C>G(Ser) +/- (1995), <b>MLH3:</b> ex1-1 c.408T>C (p=Asp) <b>MSH3:</b> ex4 c.693 G>A (p=Pro), ex20 c.2732T>G (Leu910Trp)	Left colon tubulo-villous adenoma at age 34 and a right colon adenocarcinoma at age 42; left colon moderately differentiated mucinous adenocarcinoma at age 56.	Mather dead of glioblastoma at age 63; uncle dead of metachronous colon cancer at age 50; uncle dead of bladder cancer at age 70; brother dead of glioma at age 20; brother with colon cancer at age 53 and his son with glioma at age 26; brother with colon cancer at age 36 and 53.
05/7	<b>MSH2:</b> ex5 c.942+3_+15del AAAAAAAAAA <b>PMS2:</b> ex7 c.780C>G(Ser) +/- (1995), ex13 c.2248G>A (Gly749Ser) <b>MSH3:</b> ex23 c.3131 G>a(Ala1044Thr) +/-	Transvers and descending moderately differentiated adenocarcinoma at age 34.	Father with colon cancer at age 36 and relapse at age 49 and 58; uncle dead of colon cancer at age 60; uncle dead of lung cancer at age 45 and his son with 3-4 polyps at age 42.

## Results



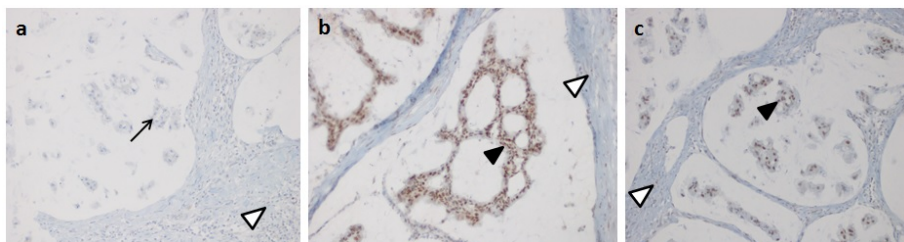
**Figure 32. (A)** Pedigree of a family with simultaneous MMR gene variants. Symbols and abbreviations used are denoted as follow: Arrows, analyzed members; black symbol, colorectal cancer or cancer associate with HNPCC; gray symbols, adenomas or cancer not associated with HNPCC; CRC, colorectal cancer; Br, brain cancer; GU, gastric ulcer; BL, bladder cancer; Bre, breast cancer; TA, tubular adenoma. Number next to diagnosis denotes age at onset; black dot, not detected. **(B)** Genotypes identified in the patients of the family; patients are identified with their pedigree number.



**Figure 33.** Microsatellite instability analysis for the II-5 family subject. The red ovals highlight the 3 unstable markers of the Bethesda panel: D2S123, BAT25 and BAT26. N: normal mucosa; T: tumor tissue.

## Results

Another interesting result of the analysis of the minor MMR genes is the absence of segregation with the disease for some pathogenetic variants identified in these genes. This condition was found in two patients carriers of the two variants identified in MSH6 gene. These are two frameshift mutations in exon 5, the c.3261\_62insC, already described in the literature, and the new mutation c.3296\_97 delTT, that alter the reading frame, thus resulting in truncated proteins (Fig. 30). The pathogenicity of one of these two variants (c.3261\_62insC) was confirmed by immunohistochemical analysis on FFPE tumor tissue sections showing the absence of the MSH6 protein compared with a lymphocyte internal control. On the other hand, MSH2 and MLH1 proteins were positive, as expected (Fig. 34).



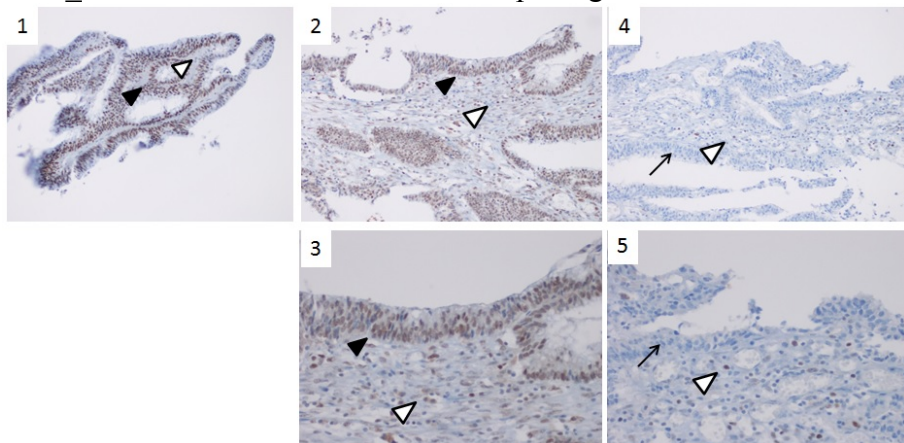
**Figure 34.** MSH6, MSH2 and MLH1 immunohistochemistry (IHC) in the patient carrier of the c.3261\_62insC variant in MSH6 exon 5. (a) Absence of MSH6 in the colon tumor cells (arrow point); (b) IHC<sup>+</sup> for MSH2 in the colon tumor cells (filled arrow head point); (c) IHC<sup>+</sup> for MLH1 in the colon tumor cells (filled arrow head point). Open arrow heads point indicate IHC<sup>+</sup> lymphocyte internal control.

The first variant was identified in a subject belonging to a family not fulfilling the Amsterdam criteria with colorectal cancers diagnosed at about the age of 60 years, while the second variant was found in a subject belonging to a family that meets the Amsterdam criteria. However, these mutations do not segregate with the disease since they were not found in the other analyzed affected relatives.

Finally, it was interesting to observe that the variant c. 2049\_2050insAGT in exon 4 of the MSH6 gene (Fig. 30) determines absence of the protein at the immunohistochemical analysis (Fig. 35), although it does not create a frameshift in the reading frame. This variant was identified in a patient with colorectal cancer diagnosed at the age of 46 years, belonging to a family that meets the Bethesda

## Results

criteria. This mutation segregates with the disease because it was found in the affected brother. Therefore, the insertion c. 2049\_2050insAGT can be considered a pathogenetic variant.



**Figure 35.** MLH1, MSH2 and MSH6 immunohistochemistry (IHC) in the patient carrier of c. 2049\_2050insAGT variant in the MSH6 exon 4. (1) IHC<sup>+</sup> for MLH1 in the colon tumor cells (filled arrow head point) 10X; (2) IHC<sup>+</sup> for MSH2 in the colon tumor cells (filled arrow head point) 10X and (3) 20X; (4) IHC<sup>-</sup> for MSH6 in the colon tumor cells (arrow point) 10X and (5) 20X. Open arrow head point indicate IHC<sup>+</sup> lymphocyte internal control.

## 4. DISCUSSION/CONCLUSIONS

The results obtained in this study provide new perspectives for the molecular diagnosis of Lynch syndrome.

Mutational analysis performed on the minor MMR genes, MSH6, PMS2, MLH3 and MSH3, was performed on 64 families with clinical diagnosis of HNPCC, selected by specific criteria on the basis of phenotypic features. We analyzed families that meet the Amsterdam I and II criteria (37%) and those that do not meet these criteria, showing an atypical phenotype (about 70%); these include familial cases with early age of onset of the disease and low microsatellite instability or cases in which the cancer developed at a young age, with a high degree of microsatellite instability. All these families were negative for point mutations and large rearrangements in two major MMR genes, namely hMLH1 and hMSH2.

Literature data report a large number of families with clinical diagnosis of Lynch syndrome, in which no mutations have been found in the major MMR genes. The Lynch-like phenotype may result from germline mutations in other genes such as the colon cancer susceptibility genes TGF $\beta$ -RII, the APC gene variants, I1307K and E1317Q, MYH, AXIN2 and EXO1. Of these, only EXO1 could have a role in the DNA repair complex, determining high microsatellite instability. Nevertheless, the literature data suggest that EXO1 is not a gene associated with Lynch syndrome (*Thompson E., 2004*). Therefore, our research was focused on the mutation detection in the minor MMR genes, assuming the low-penetrance of these genes. (*Loukola A., 2000; Yang G., 2004; Silva FC., 2009*). We carried out the mutation detection at the level of the four minor MMR genes.

Since each minor MMR gene has a repetitive sequence of 8 or more nucleotides in its coding region, they can be considered themselves as target genes of the mutator phenotype that accelerates the mutation rate in other oncogene and oncosuppressor genes.

A total of 77 variants were identified in minor MMR genes, the majority of these (about 87%) were classified variants of uncertain significance. Most of the mutations were found in MSH6 and PMS2 genes (27 and 21 variants, respectively), while a lower percentage of

variants were present in the remaining minor genes MLH3 and MSH3 (15 and 13, respectively). These results were in accordance with literature data that show mutations in the MSH6 gene in about 10% of families with Lynch syndrome and mutations in the PMS2 and MLH3 genes in about 5% of cases. In order to understand the contribution of mutations in the minor genes to the disease development, it is necessary to take into account the role of each protein in the mismatch repair complex.

Although MSH6 and MSH3 have a redundant role in the repair complex, several studies carried out in mice and in yeast demonstrate that the absence of MSH6 can not be fully compensated by the action of MSH3. Studies of Edelmann et colleagues (1997) have shown that homozygous mice for MSH6 missense mutations were unable to repair any type of defect, because the dysfunctional mutated protein was still preserved within the cell, competing with MSH3 to the MSH2 binding, and thus interfering with the MSH2-MSH3 complex to repair insertions or deletions. This suggests that missense mutations in MSH6, without the MSH3 loss, determines a severe phenotype. Instead, the loss of the MSH6 protein (nonsense mutations) associated with the severe phenotype could result from a secondary mutation in MSH3 gene. In this regard, we demonstrate that the two mutations identified in the MSH6 gene, the c.3261\_62insC and the c.3296\_97 delTT both identified in exon 5, do not segregate with the disease in the family, in the spite of their apparent pathological role due to the loss of the protein as also confirmed by the negative immunohistochemical analysis on tumor tissue. Therefore, our study was directed towards the mutation detection in other minor MMR genes PMS2, MLH3 and MSH3, that confirmed the simultaneous presence of other variants in the subject with a severe phenotype compared to other family members who showed an attenuated phenotype. Moreover, we do not exclude the presence of mutations in other genes not examined in this study, that could contribute to the disease development in synergistic manner with the MMR variants.

It is interesting to observe that the variant c. 2049\_2050insAGT in exon 4 of MSH6 gene determines the absence of the protein at the immunohistochemical analysis, although it do not create a frameshift



in the reading frame. This variant was identified in a patient who meets the Bethesda criteria, and it segregates with the disease in the family. We hypothesize that this variant causes an aberrant structural modification of the protein that induce its degradation in the cell. Therefore, the insertion c. 2049\_2050insAGT could be considered a pathogenetic variant.

Moreover, the most important result of this mutational analysis on minor MMR genes was the simultaneous presence of multiple molecular alterations in different genes or in the same gene in the DNA of a same patient. Although some MLH1 and MSH2 mutations are responsible for a monogenic transmission in tumor development, we proposed that some (or all) of the identified variants in the minor MMR genes constitute low penetrance alleles, with an additive effect on the risk of the disease (*Durauto F., 2011*). Two recent studies on the yeast genome have shown that weak alleles of the complex MMR cause a weak mutator phenotype but, when they interact with each other, a strong mutator phenotype is observed (*Martinez SL., 2010; Kumar, 2011*). This hypothesis is supported by an exhaustive molecular analysis of one of the 64 subjects analyzed in this study. In this family (Fig. 31, Results section), the simultaneous presence of several low penetrance variants was associated with a severe phenotype that could suggest an additive effect of these mutations in predisposition of cancer development.

Furthermore, we tried to define a pathogenetic role for numerous variants of uncertain significance (VUS), identified in the minor MMR genes in this study, using a multifactorial likelihood model according to literature data (*Colon cancer Family Registry 2009, InSiGHT Variant Interpretation Committee 2011*). This approach is based on both phenotypic and functional features (Introduction section). We demonstrate that none of the above criteria, including functional assays, is an indicator of pathogenicity, if considered individually; it is necessary that a combination of these strategies be used to assess an exhaustive evaluation of the pathogenicity of uncertain variants. Among these criteria, the segregation analysis should be considered the "gold standard" for the validation of any type

of in vitro analysis (*Couch FJ, 2008; SE Plon, 2008; R. Hofstra, 2008; Goldgar DE, 2008*).

The multifactorial likelihood model was also used to study other several VUS identified in major MMR genes, MLH1 and MSH2, in order to assign them a pathogenetic significance in Lynch syndrome. This multivariate approach allowed us to demonstrate a likely genotype-phenotype correlation for several variants. In this study, the analysis of 3 variants in MLH1 gene and 1 variant in MSH2 gene were reported.

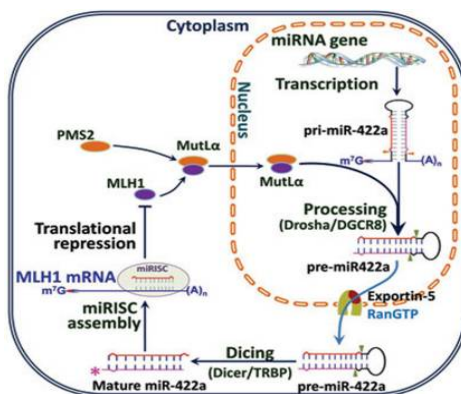
In particular, we observed that the missense mutation, c.304G>A in exon 3 of MLH1 gene, behaves as a truncating mutation at the level of the cDNA. Therefore, this can be considered as a class 5 variant and not as a class 3 variant, as reported in the literature.

Similarly, the study of the silent mutation c.438A>G in exon 5 of MLH1 gene has shown a likely role in the development of the disease as it determines aberrant effects on splicing process and consequently at the level of the final transcript.

Finally, we analyzed two variants identified in the 3'-UTR of MLH1 and MSH2 genes, according to the literature that suggests to explain the unsolved genotype-phenotype correlations in Lynch syndrome by analysing MMR gene regions not routinely investigated such as the promoter and the 5'- and the 3'-untranslated sites with respect to canonical coding regions (*Carethers JM., 2014; Boland R., 2013*).

In particular, the c.\*30\_32delTTC (3'UTR) in the MLH1 gene, yet described in literature database as class-I of uncertain significance, results in a significant reduction in both levels of the transcript and the protein. In silico analysis by computational tool TRANSFAC shows that this deletion could destroy the binding site for a transcriptional regulation factor, a glucocorticoid receptor (GR) that acts in the inflammatory response. This variant has been the subject of several literature studies that reported data in agreement with those obtained in this thesis. In particular, Mao et al. showed that the region where the mutation falls is a putative binding site for miR-422a that is disrupted by the 3-nucleotide deletion, suggesting a likely role for miR-422a in regulating expression of MLH1. The Authors demonstrated a reciprocal feedback regulation mechanism between

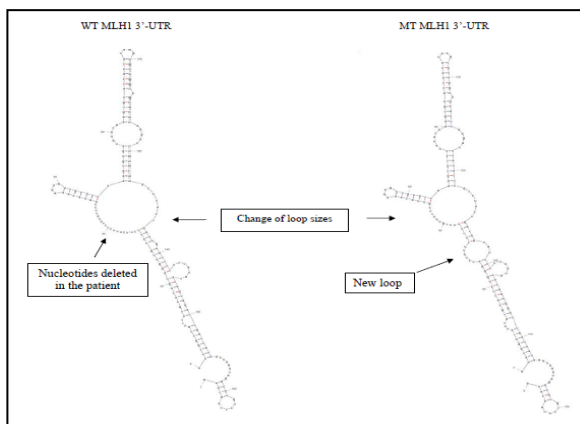
MLH1 and miR-422a where the MLH1-PMS2 heterodimer stimulate the conversion of pri-miR-422a to pre-miR-422a, while this miRNA down-regulates mutL $\alpha$  levels by suppressing MLH1 expression through base pairing with the MLH1 3'-UTR region (fig. 36) (Mao G., 2012).



**Figure 36.** Model for reciprocal feedback regulation between MLH1 and miR-422a proposed by Mao G. et al. (Cell Res 2012, 22:973-985).

A previous study highlights that this mutation confers a mutator phenotype and mismatch repair deficiency in patients with relapsed leukemia, because a significant reduction in the steady-state level of MLH1 mRNA and protein, and a consequently defective MMR activity, as demonstrated by luciferase assay, was present. The secondary structure analysis of the WT and MT 3'-UTR of MLH1 showed that the TTC-deletion lies in a large stem-loop structure. The deletion of 3-nucleotide in the region reduced the size of the stem-loop and created a new smaller stem-loop (Fig.37) (Mao G., 2008).

The alteration of this stem-loop may prevent the binding to trans-acting protein factors (TF) or microRNAs and instead may allow the interaction with different regulatory factors with a negative effect on the stability of the MLH1 transcript. In our study, this variant was always associated with a severe phenotype (MSI-H), when it was in association with the variant c.454-51t>c, described in literature as class-1 variant.



**Figure 37.** Predicted folding of MLH1 WT and MT 3'-UTR. Secondary structure of human MLH1 WT (A) or MT (B) 3'-UTR was predicted by the Mfold algorithm (<http://bioweb.pasteur.fr/seqanal/interfaces/mfold.html>). Arrows indicate changes in structures between WT and MT mRNAs (*Mao G. et al. JBC, 2008, 283: 3211-3216*)

The computational analysis performed by the HSF algorithm for this intronic variant showed a likely deleterious effect on the splicing process. Therefore, we may hypothesize a pathogenic additive effect for these two genetic alterations.

The multivariate analysis of the other 3'UTR variant, the c.\*226A>G in MSH2 gene, was also very interesting in order to understand the correlation with the disease phenotype. In our study, this variant is associated with both typical Lynch Syndrome features (CRC and endometrial atrophy) and atypical phenotype such as Hodgkin Lymphoma.

In contrast with the reduction observed with the MLH1 3'-UTR variant, the MSH2 3'UTR gene mutation determined increased mRNA and protein levels, as confirmed by immunoistochemistry analysis and by functional luciferase assay.

In silico analysis by TRANSFAC software showed this region as a putative binding site for four trans-acting protein factors (TF) with transcriptional repressor role, especially for ZNF333 (zinc-finger protein 333), and CDP (cut-like homeobox1 protein). The mutation would prevent the interaction with all these factors.

Furthermore, we identified also two potential miRNA, hsa-miR-137 and hsa-miR-4795-3p, able to bind the region where the mutation

falls, by two *in silico* prediction tools, TargetScan and miRanda. The mutation consisting in a single base substitution would prevent the binding with these above regulating factors without, however, favor the interaction with other factors. The loss of regulation for the MSH2 transcript would explain the observed overexpression that does not enhance genome stability but promotes hypermutability, as described in the Introduction section. In particular, we hypothesize that this effect is related to the loss of MSH2 down regulation by the has-miR-137 according to the literature data that indicate this miRNA as a tumor suppressor in the colon, inhibiting cell proliferation. Moreover, hsa-miR-137 is preferentially expressed in the epithelial cells of normal colonic mucosa, while low has-miR-137 expression has been observed in the adenomatous polyps and CRC tissues (Balaguer F., 2010). This condition would be enhanced by the mutation that would prevent any type of binding. Of course, this our assumption needs further study. In this study, we began to investigated the association between hsa-miR-137 and the MSH2 expression, by transfection of ectopic precursor hsa-miR-137 in two colorectal cancer cell lines, HT29 and SW480, that express high and low levels of the miR-137, respectively (Balaguer F., 2010). The qRT-PCR performed on the MSH2 mRNA after 48h from transfection showed decreased levels of the MSH2 transcript, especially in the HT29 cells, suggesting a possible role for this miRNA in MSH2 regulating expression.

In conclusion, this study allowed to clarify the genotype-phenotype correlations in Lynch syndrome, demonstrating the importance of a multifactorial likelihood analytical approach, also in order to understand the molecular mechanisms that regulate the MMR protein and their putative prognostic and therapeutic implications in Lynch Syndrome. Moreover, the simultaneous presence of molecular alterations in several genes analyzed (major and minor MMR genes) could suggest an additive effect of these mutations in the predisposition to develop cancer, in contrast with the classical monogenic transmission and in favor of a polygenic inheritance, although further studies on other familial cases are needed to confirm this hypothesis.

## 5. REFERENCES

**Alhopuro P**, Sammalkorpi H, “Candidate driver genes in microsatellite-unstable colorectal cancer”. *Int J Cancer*. 2011 May 4.

**Balaguer F**, Link A Epigenetic silencing of miR-137 is an early event in colorectal carcinogenesis. *Cancer Res*. 2010 Aug 15;70(16):6609-18.

**Bandipallian P**. “Syndrome of early onset colon cancers, hematologic malignance & features of neurofibromatosis in HNPCC families with homozygous mismatch repair gene mutations”. *Fam Cancer* 2005

**Berndt SI**, Platz EA, et al., “Mismatch repair polymorphisms and the risk of colorectal cancer”. *Int J Cancer*. 2007 Apr 1;120(7):1548-54.

**Boland CR** et al., “A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer.” *Cancer Res*. 1998 Nov 15;58(22):5248-57

**Boland CR** The Mystery of Mismatch Repair Deficiency: Lynch or Lynch-like?. *Gastroenterology* 2013;144:868–881.

**Buchanan DD**, Christophe Rosty, Mark Clendenning et al. Clinical problems of colorectal cancer and endometrial cancer cases with unknown cause of tumor mismatch repair deficiency (suspected Lynch syndrome). *The Application of Clinical Genetics* 2014;7 183–193.

**Cannavo E**, et al., “Expression of the MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res*. 2005 Dec 1;65(23):10759-66.

**Carethers JM**. Differentiating Lynch-Like From Lynch Syndrome. *Gastroenterology* 2014;146:602–614.

**Chao E. C.**, Lipkin S. M. “Molecular models for the tissue specificity of DNA mismatch repair-deficient carcinogenesis”. *Nucleic Acids Res*. 2006 Jan

**Chung CC**, Chanock SJ. Current status of genome-wide association studies in cancer. *Hum Genet*. 2011 Jul;130(1):59-78

**Corso G**, Marrelli D, Roviello F. Familial gastric cancer: update for practice management *Fam Cancer*. 2011 Jun;10(2):391-6.

**Couch FJ.**, Rasmussen LJ, et al., “Assessment of functional effects of unclassified genetic variants”. *Hum Mutat*. 2008 Nov;29(11):1314-26.

## References

---

- Dong Y**, Yu J, Ng SS. MicroRNA dysregulation as a prognostic biomarker in colorectal cancer. *Cancer Manag Res*. 2014 Oct 14;6:405-22
- Dragileva E**, Hendricks A., et al., "Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes". *Neurobiol Dis*. 2009 Jan;33(1):37-47.
- Durauto F**, Liccardo R, et al., "Association of low-risk MSH3 and MSH2 variant alleles with Lynch syndrome: probability of synergistic effects". *Int J Cancer*. 2011 Oct 1;129(7):1643-50.
- Durauto F**, R. Liccardo, A. Cavallo, M. De Rosa and P. Izzo. Synergistic Effects of Low-Risk Variant Alleles in Cancer Predisposition. *Carcinogenesis*, book edited by Kathryn Tonissen, ISBN 978-953-51-0945-7, Published: January 23, 2013
- Edelmann W**, Yang K, et al. "Mutation in the mismatch repair gene MSH6 causes cancer susceptibility". *Cell*. 1997 Nov 14;91(4):467-77.
- Elrasheid A. H.** Kheirelseid, Nicola Miller, Kah Hoong Chang, Catherine Curran et al. Mismatch repair protein expression in colorectal cancer. *J Gastrointest Oncol* 2013;4(4):397-408.
- Gylling A**, Abdel-Rahman WM, et al Is gastric cancer part of the tumour spectrum of hereditary non-polyposis colorectal cancer? A molecular genetic study. *Gut*. 2007 Jul;56(7):926-33.
- Goldgar DE.**, Easton DF., et al., "Genetic evidence and integration of various data sources for classifying uncertain variants into a single model". *Hum Mutat*. 2008 Nov;29(11):1265-72.
- Haugen et al.**, "Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer". *Cancer Res*. 2008 Oct 15;68(20):8465-72.
- Herkert JC**, Niessen RC., et al., "Paediatric intestinal cancer and polyposis due to bi-allelic PMS2 mutations: case series, review and follow-up guidelines". *Eur J Cancer*. 2011 May;47(7):965-82.
- Hofstra RM**, Spurdle AB, et al., "Tumor characteristics as an analytic tool for classifying genetic variants of uncertain clinical significance". *Hum Mutat*. 2008 Nov;29(11):1292-303.
- Hsieh P**, Yamane K. "DNA mismatch repair: molecular mechanism, cancer, and ageing". *Mech Ageing Dev*. 2008 Jul-Aug.
- Huang J**, Kuismanen SA, et al. "MSH6 and MSH3 are rarely involved in genetic predisposition to nonpolyptotic colon cancer". *Cancer Res*. 2001 Feb 15;61(4):1619-23.
- Humar A**, Boland CR, et al. "Revised Bethesda Guidelines for hereditary non polyposis colorectal cancer (Lynce Syndrome) and microsatellite instability". *Journal Natl Cancer Inst*. 2004 Feb; 18;96(4):261-8.

## *References*

---

- Imai K**, Yamamoto H. “Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics”. *Carcinogenesis* 2008 Apr;29(4):673-80.
- Ito E**, et al A core promoter and a frequent single-nucleotide polymorphism of the mismatch repair gene hMLH1. *Biochem Biophys Res Commun.* 1999 Mar 24;256(3):488-94.
- Iwahashi Y**, Ito E, Yanagisawa Y, et al. Promoter analysis of the human mismatch repair gene hMSH2. *Gene.* 1998 Jun 15;213(1-2):141-7.
- Jelsig AM**, Niels Qvist, Klaus Brusgaard, Claus Buhl Nielsen et al. Hamartomatous polyposis syndromes: A review. *Orphanet Journal of Rare Diseases* 2014, 9:101.
- Ji G**, Long Y, Zhou Y, Huang C, Gu A, Wang X. Common variants in mismatch repair genes associated with increased risk of sperm DNA damage and male infertility. *BMC Med.* 2012 May 17;10:49.
- Jiang C**, Zhao ML, Waters KM, Diaz M. Activation-induced deaminase contributes to the antibody-independent role of B cells in the development of autoimmunity. *Autoimmunity.* 2012 Sep;45(6):440-8..
- Jianghua Ou**, Renée C. Niessen, Anne Lützen, Rolf H. Sijmons, Jan. H. Kleibeuker et al. Functional Analysis Helps to Clarify the Clinical Importance of Unclassified Variants in DNA Mismatch Repair Genes. *Hum Mut* 2007 28(11),1047\_1054.
- Jiricny J**. “The multifaceted mismatch-repair system”. *Mol Cell Biol* 2006 May.
- Jun S.-H.**, et al., “DNA mismatch repair system. Classical and fresh roles.” *FEBS Journal* 2006 Feb10; 273: 1609-19.
- Kheirleaid EA**, Miller N, Chang KH, Curran C, Hennessey E, Sheehan M, Kerin MJ. Mismatch repair protein expression in colorectal cancer. *J Gastrointest Oncol.* 2013 Dec;4(4):397-408.
- Kang SY**, Park CK, Chang DK, et al Lynch-like syndrome: Characterization and comparison with EPCAM deletion carriers. *Int J Cancer.* 2015 Apr;136(7):1568-78.
- Kim JO**, Gyeong Hoon Kang. Molecular and prognostic heterogeneity of microsatellite-unstable colorectal cancer. *World J Gastroenterol* 2014 April 21; 20(15): 4230-4243.
- Knudson A.G**. “Hereditary cancer:two hits revisited”. *J Cancer Res Clin Oncol* 1996. 122(3):135-40.
- Koornstra J**. et al., “Management of extracolonic tumours in patients with Lynch syndrome”. *Lancet Oncol.* 2009;10:400-08.



## References

---

- Kumar C.**, Piacente SC, et al., “Multiple factors insulate Msh2-Msh6 mismatch repair activity from defects in Msh2 domain I”. *J Mol Biol.* 2011 Aug 26;411(4):765-80.
- Landau DA** and Frank J. Slack. MicroRNAs in Mutagenesis, Genomic Instability and DNA Repair. *Semin Oncol.* 2011 December ; 38(6): 743–751. doi:10.1053/j.seminoncol.2011.08.003.
- Lanza G**, R. Gafà, A. Santini, et al Immunohistochemical Test for MLH1 and MSH2 Expression Predicts Clinical Outcome in Stage II and III Colorectal Cancer Patients. *J Clin Oncol* 2006 24:2359-2367.
- Lawes DA**, SenGupta SB, et al. “Pathogenesis and clinical management of hereditary non-polyposis colorectal cancer”. *Surg.* 2002 Nov; 89(11):1357-69.
- Lee SY**, Chung H., “Microsatellite alterations at selected tetranucleotide repeats are associated with morphologies of colorectal neoplasias”. *Gastroenterology* 2010 Nov;139(5):1519-25.
- Li Z.**, Peled J. U., et al., “A role for Mlh3 in somatic hypermutation”. *DNA Rep.* 2006 Mar 27; 5: 675-82.
- Lindor NM**, et al. “Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency: familial colorectal cancer type X”. *JAMA.* 2005 Apr 27;293(16):1979-85.
- Liu H.-X.**, et al., “The Role of hMLH3 in Familial Colorectal Cancer”. *Cancer Res.* 2003 Apr 15;63: 1894-99.
- Loukola A.**, et al., “Germline and Somatic Mutation Analysis of MLH3 in MSI-Positive Colorectal Cancer”. *Am. J Pathol.* 2000 Aug. 2; 157(2): 347-52.
- Lucci-Cordisco E**, Rovella V, et al. "Mutations of the 'minor' mismatch repair gene MSH6 in typical and atypical hereditary nonpolyposis colorectal cancer". *Fam Cancer.* 2001;1(2):93-9.
- Lynch HT**, Kristen Drescher, Joseph Knezetic, Stephen Lanspa. Genetics, Biomarkers, Hereditary Cancer Syndrome Diagnosis, Heterogeneity and Treatment: A Review. *Current Treatment Options in Oncology* (2014) 15:429–442.
- Mao G**, Pan X, Gu L. Evidence that a mutation in the MLH1 3'-untranslated region confers a mutator phenotype and mismatch repair deficiency in patients with relapsed leukemia. *J Biol Chem.* 2008 Feb 8;283(6):3211-6
- Mao G**, Lee S, Ortega J, Gu L, Li GM. Modulation of microRNA processing by mismatch repair protein MutL $\alpha$ . *Cell Res.* 2012 Jun;22(6):973-85

## References

---

- Marra G**, Iaccarino I, et al Mismatch repair deficiency associated with overexpression of the MSH3 gene. *Proc Natl Acad Sci U S A*. 1998 Jul 21;95(15):8568-73.
- Martinez SL**, Kolodner RD., “Functional analysis of human mismatch repair gene mutations identifies weak alleles and polymorphisms capable of polygenic interactions”. *Proc Natl Acad Sci U S A*. 2010 Mar 16;107(11):5070-5.
- Mensenkamp AR**, Vogelaar IP, et al Somatic mutations in MLH1 and MSH2 are a frequent cause of mismatch-repair deficiency in Lynch syndrome-like tumors. *Gastroenterology*. 2014 Mar;146(3):643-646
- Mrkonjic M**, Stavroula Raptis, Roger C.Green, Neerav Monga, Darshana Daftary et al. MSH2 - 118T>C and MSH6 - 159C>T promoter polymorphisms and the risk of colorectal cancer. *Carcinogenesis* vol.28 no.12 pp.2575–2580, 2007.
- Nicholson A.**, Hendrix M., et al., “Regulation of Mitotic Homeologous Recombination in Yeast: Functions of Mismatch Repair and Nucleotide Excision Repair Genes”. *Genetics* 2000 Jan; 154: 133-46.
- O’Brien V.**, Brown R. “Signalling cell cycle and cell death through the MMR System”. *Carcinogenesis* 2006; 27(4): 682-94.
- Olschwang S.**, Bonaiti C., et al., “HNPCC syndrome (Hereditary Non Polyposis Colon Cancer) : identification and management”. *RevMed* 2006 Jan May;54(4):215-29.
- Penã-Diaz J**, Stephanie Bregenhorn, Medini Ghodgaonkar, Cindy Follonier et al. Noncanonical Mismatch Repair as a Source of Genomic Instability in Human Cells. *Molecular Cell* 2012 Sep 14;47(5):669-80
- Peltomaki P**. “DNA mismatch repair and cancer”. *Mutat Res*. 2001 Mar; 488(1):77-85.
- Peters A**, Born H, et al., “Compound heterozygosity for MSH6 mutations in a pediatric lymphoma patient”. *J Pediatr Hematol Oncol*. 2009 Feb;31(2):113-5.
- Plaschke J**, Kruger S, et al. "Eight novel MSH6 germline mutations in patients with familial and nonfamilial colorectal cancer selected by loss of protein expression in tumor tissue". *Hum Mutat*. 2004 Mar;23(3):285.
- Plon SE.**, Eccles DM., et al., “Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results”. *Hum Mutat*. 2008 Nov;29(11):1282-91.
- Poley JW**, Wagner A., et al., “Biallelic germline mutations of mismatch-repair genes: a possible cause for multiple pediatric malignancies”. *Cancer*. 2007 Jun 1;109(11):2349-56.

## References

---

- Raptis S**, Mrkonjic M et al MLH1 -93G>A promoter polymorphism and the risk of microsatellite-unstable colorectal cancer. *J Natl Cancer Inst.* 2007 Mar 21;99(6):463-74.
- Roa S**, Li Z, et al., “MSH2/MSH6 complex promotes error-free repair of AID-induced dU:G mispairs as well as error-prone hypermutation of A:T sites”. *PLoS One* 2010 Jun 17;5(6):e11182.
- Samadder NJ**, Jasperson K, Burt RW. Hereditary and Common Familial Colorectal Cancer: Evidence for Colorectal Screening. *Dig Dis Sci.* 2014 Dec 12
- Sargent DJ**, Silvia Marsoni, Genevieve Monges, Stephen N. Thibodeau et al. Defective Mismatch Repair As a Predictive Marker for Lack of Efficacy of Fluorouracil-Based Adjuvant Therapy in Colon Cancer. *Journal Of Clinical Oncology* Volume 28, Number 20, July 10 2010.
- Seriola A**, Spits C., et al., “Huntington's and myotonic dystrophy hESCs: down-regulated trinucleotide repeat instability and mismatch repair machinery expression upon differentiation”. *Hum Mol Genet.* 2011 Jan 1;20(1):176-85.
- Sharma et al.** “BRAF mutation testing in colorectal cancer”. *Arch Pathol Lab Med.* 2010 Aug;134(8):1225-8.
- Shcherbakova PV**, Hall MC, et al Inactivation of DNA mismatch repair by increased expression of yeast MLH1. *Mol Cell Biol.* 2001 Feb;21(3):940-51.
- Sinicrope FA**, Sargent DJ. Molecular pathways: microsatellite instability in colorectal cancer: prognostic, predictive, and therapeutic implications. *Clin Cancer Res.* 2012 Mar 15;18(6):1506-12
- Silva FC**, et al. “Mismatch repair genes in Lynch syndrome: a review”. *Sao Paulo Med J.* 2009 Jan 127(1):46-51.
- Smith TR.**, Liu-Mares W., et al. “ Genetic polymorphisms of multiple DNA repair pathways impact age at diagnosis and TP53 mutations in breast cancer”. *Carcinogenesis.* 2011 Sep;32(9):1354-60.
- Stoffel EM**, Pamela B. Mangu, Stephen B. Gruber et al. Hereditary Colorectal Cancer Syndromes: American Society of Clinical Oncology Clinical Practice Guideline Endorsement of the Familial Risk–Colorectal Cancer: European Society for Medical Oncology Clinical Practice Guidelines. *J Clin Oncol* 2015 33:209-217.
- Suraweera N.** et al., “Evaluation of tumor microsatellite instability using five quasimonomorphic nucleotide repeats and pentaplex PCR “. *Gastroenterology* 2002. Dec;123(6):1804-11.

## References

---

**Svrcek M**, El-Murr N, Overexpression of microRNAs-155 and 21 targeting mismatch repair proteins in inflammatory bowel diseases. *Carcinogenesis*. 2013 Apr;34(4):828-34

**Syngal S**, Fox EA, et al., "Interpretation of genetic test results for hereditary nonpolyposis colorectal cancer: implications for clinical predisposition testing". *JAMA*. 1999 Jul 21;282(3):247-53.

**Thompson E**, Meldrum CJ, et al., "Hereditary non-polyposis colorectal cancer and the role of hPMS2 and hEXO1 mutations". *Clin Genet*. 2004 Mar;65(3):215-25.

**Tomé S**, Holt I, et al., "MSH2 ATPase domain mutation affects CTG\*CAG repeat instability in transgenic mice". *PLoS Genet*. 2009 May;5(5):e1000482.

**Umar A**, Risinger JJ, et al., "Functional overlap in mismatch repair by human MSH3 and MSH6". *Genetics*. 1998 Apr;148(4):1637-46.

**Valle L**. Genetic predisposition to colorectal cancer: Where we stand and future perspectives. *World J Gastroenterol* 2014 August 7; 20(29): 9828-9849.

**Valeri N**, Gasparini P, et al MicroRNA-21 induces resistance to 5-fluorouracil by down-regulating human DNA MutS homolog 2 (hMSH2). *Proc Natl Acad Sci U S A*. 2010 Dec 7;107(49):21098-103.

**van Lier MGF**, Anja Wagner, Monique E. van Leerdam et al. A review on the molecular diagnostics of Lynch syndrome: a central role for the pathology laboratory. *J. Cell. Mol. Med*. Vol 14, No 1-2, 2010 pp. 181-197.

**Vilar E**, Maureen E. Mork, Amanda Cuddy, Ester Borrás et al. Role of microsatellite instability-low as a diagnostic biomarker of Lynch syndrome in colorectal cancer. *Cancer Genetics* 207 (2014) 495-502.

**Wang L**, Cunningham JM, Winters JL, et al BRAF mutations in colon cancer are not likely attributable to defective DNA mismatch repair. *Cancer Res*. 2003 Sep 1;63(17):5209-12.

**Wang Y**, Friedl W., "Hereditary nonpolyposis colorectal cancer: causative role of a germline missense mutation in the *hMLH1* gene confirmed by the independent occurrence of the same somatic mutation in tumour tissue". *Hum Genet*. 1997 Sep;100(3-4):362-4.

**Wang Y**, Taniguchi T. MicroRNAs and DNA damage response: implications for cancer therapy *Cell Cycle*. 2013 Jan 1;12(1):32-42.

**Weitz J**, Koch M, et al. "Colorectal cancer". *Lancet*. 2005 Jan 8-14;365(9454):153-65.

**Wiesendanger M**, et al., "Somatic Hypermethylation in MutS Homologue MSH3-, MSH6-, and MSH3/MSH6-deficient Mice Reveals a Role for the MSH2-MSH6

## References

---

Heterodimer in Modulating the Base substitution Pattern". J Exp. Med. 2000 Feb 7; 191(3):579-584.

**Xicola RM**, Llor X, Pons E, et al; Performance of different microsatellite marker panels for detection of mismatch repair-deficient colorectal tumors Gastrointestinal Oncology Group of the Spanish Gastroenterological Association. J Natl Cancer Inst. 2007 Feb 7;99(3):244-52.

**Yamane K**, et al. "BRCA1 activates a G2-M cell cycle checkpoint following 6-thioguanin-induced DNA mismatch damage". Cancer Res. 2007 Jul 1;67(13):6286-92.

**Yang G.**, Scherer SJ, et al. " Dominant effects of an Msh6 missense mutation on DNA repair and cancer susceptibility". Cancer Cell. 2004 Aug;6(2):139-50.

**Yemin Wang** and Toshiyasu Taniguchi. MicroRNAs and DNA damage response. Cell Cycle 12:1, 32–42; January 1, 2013.

**Yujuan D**, Jun Yu, Simon SM Ng. MicroRNA dysregulation as a prognostic biomarker in colorectal cancer. Cancer Management and Research 14 October 2014

**Zhang G**, Studies on the interactions between human replication factor C and human proliferating cell nuclear antigen Proc Natl Acad Sci U S A. 1999 Mar 2;96(5):1869-74.

**Zaanan A**, Meunier K, et al Microsatellite instability in colorectal cancer: from molecular oncogenic mechanisms to clinical implications Cell Oncol (Dordr). 2011 Jun;34(3):155-76

**Zhang J**, Lindroos A. "Gene conversion is a frequent mechanism of inactivation of the wild-type allele in cancer from MLH1/MSH2 deletion carriers". Cancer Res 2006.

**Zhang J.**, Chiodin R Badra A., Zhang G. The impact of next-generation sequencing on genomics. J Genet Genomics 2011; 38(3): 95-109.

<http://blocks.fhcrc.org/sift/SIFT.html>

<http://genetics.bwh.harvard.edu/pph/>

<https://www.predictprotein.org/>

[www.umd.be/HSF/](http://www.umd.be/HSF/)

[www.ncbi.nlm.nih.gov/tools/primer-blast](http://www.ncbi.nlm.nih.gov/tools/primer-blast)

<http://www.biobaseinternational.com/product/transcription-factor-binding-sites>

<http://www.targetscan.org>, <http://www.microrna.org>, <http://mirdb.org>

[www.insight-group.org](http://www.insight-group.org)

[www.mrc-holland.nd](http://www.mrc-holland.nd).